

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.

PCTWORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁶ : C12N	A2	(11) International Publication Number: WO 97/47732 (43) International Publication Date: 18 December 1997 (18.12.97)
(21) International Application Number: PCT/US97/12599 (22) International Filing Date: 13 June 1997 (13.06.97) (30) Priority Data: 60/019,934 14 June 1996 (14.06.96) US 60/043,948 19 February 1997 (19.02.97) US 60/038,915 21 February 1997 (21.02.97) US (71) Applicant (for all designated States except US): SMITHKLINE BEECHAM CORPORATION [US/US]; One Franklin Plaza, Philadelphia, PA 19103 (US). (72) Inventors; and (73) Inventors/Applicants (for US only): CHAIKIN, Margery, Ann [US/US]; 242 Carriage Court, Wayne, PA 19087 (US). LYN, Sally, Doreen, Patricia [CA/US]; 134-B Hampton Court, West Chester, PA 19380 (US). SWEET, Raymond, Whitney [US/US]; 108 Edgehill Road, Bala Cynwyd, PA 19004 (US). TRUNEH, Alemseged [US/US]; 1008 Stoneham Drive, West Chester, PA 19382 (US). (74) Agents: GIMMI, Edward, R. et al.; SmithKline Beecham Corporation, Corporate Intellectual Property, UW2220, 709 Swedeland Road, P.O. Box 1539, King of Prussia, PA 19406-0939 (US).		(81) Designated States: AL, AM, AU, BB, BG, BR, CA, CN, CZ, EE, GE, GH, HU, IL, IS, JP, KG, KP, KR, LK, LR, LT, LV, MD, MG, MK, MN, MX, NO, NZ, PL, RO, SG, SI, SK, TR, TT, UA, US, UZ, VN, YU, ARIPO patent (GH, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG). Published <i>Without international search report and to be republished upon receipt of that report.</i>
(54) Title: HEXAMERIC FUSION PROTEINS AND USES THEREFOR		
(57) Abstract A hexameric fusion protein containing a dimeric binding protein provided with a tailpiece from an IgA antibody is described. This fusion protein is useful in therapeutics and vaccines, but is particularly well suited for applications for which the binding protein from which it is derived is unsatisfactory because of low binding affinity or for applications where multivalency is desired. These applications include diagnostics, binding assays and screening assays.		

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NR	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakhstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

HEXAMERIC FUSION PROTEINS AND USES THEREFOR

Related Application

This application claims benefit of U.S. Provisional patent applications numbered
5 60/019,934 filed June 14, 1996, 60/043,948 filed February 19, 1997 and 60/038,915 filed
February 21, 1997.

Background of the Invention

IgM and IgA are the two classes of human antibodies that form homo-oligomeric
10 structures. By far the most extensively studied of these is IgM.

The classical view of IgM structure is as a pentamer in combination with a single
copy of a second protein, the J-chain that becomes associated with IgM during its assembly
and export. This J-chain can covalently associate with IgM through the formation of a
disulfide bond between a cysteine residue in the J chain and a cysteine residue in a short 18
15 amino acid extension, designated μ tp, from the canonical C-terminal constant region of the
heavy chain. This cysteine residue appears to be required for the formation of the IgM
pentamer in association with the J-chain [A. C. Davis *et al.*, EMBO J., 8(9):2519-2526
(1989)]. However, a hexameric form of IgM, devoid of the J-chain, was described and
characterized over two decades ago and has recently been characterized in more detail in
20 terms of its biochemical and potential biological activities [reviewed in Brewer *et al.*,
Immunology Today, 15:165-168 (1994)]. The production of oligomeric IgG proteins has
been achieved by addition of the 18 amino acid IgM tailpiece segment (μ tp) to the α tp
corresponding C-termini end of the C γ 3 region of the IgG1-4 proteins by DNA recombinant
technology [R. I. F. Smith and S. L. Morrison, Biotechnology, 12:683-688 (1994); R. I. F.
25 Smith *et al.*, J. Immunol., 154:2226-2236 (1995)].

Human IgA also has an 18 amino acid tailpiece segment (α tp) which bears some
sequence homology to μ tp. In man, there are two α constant region loci which encode distinct
sequences, but the tailpiece regions for the α 1 and α 2 regions are quite similar, or in some
cases reported to be identical [*Sequences of Proteins of Immunological Interest*, fifth edition,
30 EA Kabat *et al.*, Vol. 1, U.S. Department of Health and Human Services, NIH publication
no. 91-3242, (1991)]. However, unlike IgM, IgA occurs most frequently as a monomer
antibody, similar to the IgG subclasses, or as a dimer antibody plus one molecule of J-chain
[Mestecky and Kilian, Methods in Enzymology, 116:37-75 (1985); T. B. Tomasi, Immun.

Today, 13:416-418 (1992)]. Higher oligomers/aggregates of IgA are reported [Mestecky and Kilian, cited above], but these are poorly characterized components in complex mixtures containing other proteins interactive with IgA. Recombinant IgA has been expressed in the presence and absence of the J chain (Bruggemann *et al.*, J. Exp. Med., 166:1351-1361 (1987); Morton *et al.*, J. Immunol., 151:4743-4752 (1993); Carayannopoulos *et al.*, Proc. Natl Acad Sci. USA, 91:8348-8352 (1994); Terskikh *et al.*, Mol. Immunol., 31:1313-1319 (1994)]. The IgA proteins produced in the absence of the J chain were monomeric or dimeric forms by nonreducing SDS/PAGE and appeared as dimers in solution. In one study (Carayannopoulos *et al.*, above), the co-expression of the J-chain led to formation of disulfide linked IgA dimers together with J chain.

The CD28 receptor, a member of the immunoglobulin superfamily of molecules (IgSF) [A.F. Williams and A.N. Barclay, Annu. Rev. Immunol., 6:381-405 (1988)], is a 44 kDa homodimer glycoprotein expressed on the surface of T-lineage cells including thymocytes and peripheral T cells in the spleen, lymph node and peripheral blood. CD28 interacts with two different counter-receptors CD80 (also known as B7 and B7.1) [P. S. Linsley *et al.*, Proc. Natl. Acad. Sci. USA, 87(13):5031-5035 (1990); G. J. Freeman *et al.*, J. Exp. Med., 174(3):625-631 (1991)] and CD86 (also called B7.2 and B70) [M. Azuma *et al.*, Nature, 366(6450):76-79 (1993); G. J. Freeman *et al.*, J. Exp. Med., 178(6):2185-2192 (1993); G.J. Freeman *et al.*, Science, 262(5135):909-911 (1993)], expressed on antigen presenting cells (APCs), to deliver crucial co-stimulatory signals for sustained activation of T cells, through its association via the cytoplasmic domain with PI3-kinase [F. Pages *et al.*, Nature, 369(6478):327-329 (1994); P. H. Stein *et al.*, Molecul. & Cell. Biol., 14(5):3392-3402 (1994)] and other signalling pathways [K.E. Truitt *et al.*, J. Immunol., 155:4702-4710 (1995); J. A. Nunes *et al.*, J. Biol. Chem., 271(3):1591-1598 (1996); H. Schweider *et al.*, Eur. J. Immunol., 25:1044-1050 (1995)]. Both CD80 [P. S. Linsley *et al.*, J. Exp. Med., 174(3):561-569 (1991)] and CD86 [Azuma *et al.*, cited above; Freeman *et al.*, 1993, cited above; Freeman *et al.*, 1993, cited above] also recognize CTLA-4 [J.F. Brunet *et al.*, Nature, 328(6127):267-270 (1987)], a homolog of CD28, expressed transiently and at low receptor density on activated CD8⁺ and CD4⁺ T cells.

Antagonism of CD28 interactions with the CD80 or CD86 counter-receptors using CTLA4-Ig fusion proteins or antibodies directed against CD80 and CD86 inhibits T cell activation *in vitro*, suppresses humoral and cellular immune responses *in vivo*, inhibits graft rejection and the progression of autoimmune diseases *in vivo* [reviewed in J. A. Bluestone,

Immunity, 2:555-559 (1995); Harlan *et al.*, Clin. Immunol. and Immunopath., 75(2):99-111 (1995)]. Thus, CD28 is a target for development of immunosuppressive agents. To identify small molecule antagonists, a rapid and reproducible assay is desirable for the screening of synthetic compounds, natural products, and peptides. Particularly desirable is a protein
5 based assay which would isolate the receptor and its counter-receptor from interference by other components of cell-based assays, and which is additionally adaptable to automation. The affinity of the interaction of CD28 with both counter receptors is quite low [P. S. Linsley *et al.*, Immunity, 1:793-801 (1994)], with an approximate K_d of 200 nM for the binding of a soluble CD80-Ig fusion protein to an immobilized CD28-Ig fusion protein [P. S. Linsley *et al.*, J. Exp. Med., 173(3):721-730 (1991)]. This low affinity hampers development of a
10 sensitive protein binding assays amenable to screening many compounds.

What is needed is a method for increasing the avidity of binding proteins, particularly those with low affinity, for use in screening and diagnostic assays, therapeutics, and vaccines.

15 Summary of the Invention

In one aspect, the present invention provides a hexameric fusion protein which provides increased binding activity as compared to the protein from which it is derived and methods of making same. This fusion protein is particularly useful in binding assays and may be readily purified.

20 The hexameric fusion protein of the invention contains a dimeric binding protein and a tailpiece (α tp) characterized by the activity of the tailpiece from the C-terminus of the heavy chain of an IgA antibody. In one embodiment, the binding protein is a natively dimeric binding protein or a functional fragment thereof. In another embodiment, the binding protein is recombinantly engineered to have a dimeric form. This is preferably achieved by fusion of
25 a protein fragment which contains the extracellular domain of a selected binding protein to an Fc fragment. These binding proteins, when provided with the α tp, assemble into homo- or hetero-hexamers.

In yet another aspect, the present invention provides a polynucleotide sequence encoding a stable hexameric fusion protein of the invention.

30 In a further aspect, the present invention provides a vector comprising the above-described polynucleotide sequence and a sequence controlling expression of the fusion protein in a selected host cell.

In still another aspect, the present invention provides a recombinant host cell containing the above-described vector.

In a further aspect, the present invention provides methods of producing and purifying a stable hexameric fusion protein by providing a host cell containing the stable
5 hexameric fusion protein of the invention, recovering the stable hexameric fusion protein, and purifying the recovered protein. The strands of the fusion protein are preferably co-produced and assembled in the host cell.

In still a further aspect, the present invention provides a pharmaceutical composition containing a stable hexameric fusion protein or a DNA sequence encoding the stable
10 hexameric fusion protein of the invention and a pharmaceutically acceptable carrier.

In yet another aspect, the present invention provides for screening for ligands to a hexameric fusion protein of the invention. Also provided are assays for inhibitors of hexameric binding protein/ligand interaction.

Other aspects and advantages of the present invention are described further in the
15 following detailed description of the preferred embodiments thereof.

Brief Description of the Drawings

Fig. 1 is a schematic representation of the hexameric CD80-Ig α p protein of the invention. The regions of the molecule corresponding to the CD80 extracellular domain, the
20 IgG1 hinge, CH2, and CH3 domains, and the α p segment are indicated. The letter "S" in the diagram indicates the positions of predicted disulfide bonds between cysteine residues.

Fig. 2 is a plasmid map illustrating the expression construct for the CD80-Ig α p protein of the invention. The plasmid is 7,167 base pairs in size. Beginning at residue 1 in a clockwise manner: "cmv pro" is the major late CMV promoter for transcription of the
25 downstream CD80-Ig α p coding sequence; "CD80" encodes the signal peptide and extracellular domain of human CD80; "Fc" encodes the hinge, CH2, and CH3 regions of human IgG1; " α p" encodes the human α p segment; "BGH" is the polyadenylation signal region from the bovine growth hormone gene; "betaglobin" is the mouse major b-globin promoter; "dhfr" encodes the mouse dhfr(dihydrofolate reductase) protein; "SV40" is the
30 SV40 early polyadenylation region; and "ori" and "amp" are the bacterial origin of replication and beta lactamase gene, respectively, from the common cloning plasmid pBR322. The corresponding plasmids CD86Fc α p link and CTLA4Fc α p link were constructed for the expression of the CD86-Ig α p and CTLA4-Ig α p proteins (see Figs. 5 and 6).

Fig. 3A-3H is the complete DNA sequence of the CD80Fc α tp link plasmid [SEQ ID NO: 1] shown in Fig. 2.

Fig. 4A-4D is the DNA and encoded protein [SEQ ID NOS: 2 and 3] sequences for the CD80-Ig α tp region in the vector CD80-Fc α tp link. Bolded regions show restriction sites for reference to Fig. 2 and the initiation codon, mature processing site, hinge region, and C-terminal α tp segment.

Fig. 5A-5B is the DNA and encoded protein sequences [SEQ ID NOS: 4 and 5] for the extracellular domain of CD86 in the vector CD86Fc α tp link. The sequence outside of the Kpn I and Eag I sites is the same as for CD80Fc α tp link (see Figs. 3A-3H and 4A-4D).

Fig. 6A-6C is the DNA and encoded protein sequences [SEQ ID NOS: 6 and 7] for the CMV promoter and the extracellular domain of CTLA-4 in the vector CTLA4-Fc α tp link. The sequence 5' to base 514 and 3' of the Eag I site is the same as for CD80Fc α tp link.

Fig. 7 is a profile for chromatography of CD80-Ig α tp on a Superdex 200 column. The first peak eluting at about 45 min is the hexameric protein complex while the second peak migrates at the position observed for monomeric CD80-Ig. The inset shows a coomassie stained pattern for the purified CD80-Ig α tp protein on SDS/PAGE under reducing (R) and nonreducing (NR) conditions.

Fig. 8 is a chart showing equilibrium sedimentation (main panel) and sedimentation velocity (inset) analytical centrifugation of the CD80-Ig α tp protein with a modeled fit to a hexamer/(hexamer)₂ equilibrium. The upper graph shows the residuals for the equilibrium sedimentation centrifugation.

Fig. 9 is a line graph illustrating the binding of biotinylated CD80-Ig α tp (labeled B7-FcA) to CD28-Ig immobilized at three different concentrations in an ELISA format. Binding was inhibited by the mAb CD28.1 or by CTLA4-Ig.

Fig. 10 is a line graph illustrating the binding of biotinylated CD80-Ig α tp, CD86-Ig α tp, and CD80-Ig compared to immobilized CD28-Ig in an ELISA format.

Fig. 11 is a line graph illustrating the binding of biotinylated CD80-Ig α tp, CD86-Ig α tp, and CD80-Ig compared to immobilized CTLA4-Ig in an ELISA format.

Fig. 12 is a line graph illustrating the competition of biotinylated CD80-Ig α tp binding to immobilized CD28-Ig (coated at 200 mg/ml) by CD80-Ig α tp itself, CD80-Ig, CTLA4-Ig, and CD28.2 MAb.

Fig. 13A is a line graph illustrating the binding of CD80-Ig α tp to wild-type and mutant immobilized CD28-muIg2a proteins.

Fig. 13B is a line graph illustrating the binding of CD86-Ig α tp to wild-type and mutant immobilized CD28-muIg2a proteins.

5 Fig. 13C is a line graph illustrating the binding of rabbit polyclonal antisera to wild-type and mutant immobilized CD28-muIg2a proteins.

Fig. 14 is a chart illustrating sequentially the binding of CD80-Ig and CD80-Ig α tp to CD28-Ig immobilized on a biosensor chip as measured by surface plasmon resonance.

10 Fig. 15 is a chart illustrating the binding of CD80-Ig α tp and CD86-Ig α tp to CD28-Ig immobilized on a biosensor chip as measured by surface plasmon resonance.

Figs. 16A and 16B are line graphs illustrating the binding of CD80-Ig α tp and CD86-Ig α tp, respectively, to cells expressing human CD28 on their surface in the presence or absence of a CD28 monoclonal antibody that inhibits this interaction.

15 Fig. 17 is a bar chart illustrating the level of IL-2 production by PCD28.1 cells treated with monomeric and hexameric CD80 (labeled B7.1-Ig and B7.1-IgA, respectively) and CD86 (labeled B7.2-Ig and B7.2-IgA, respectively) Ig fusion proteins. The proteins were used (1) alone in solution, (2) alone immobilized through goat anti-human antibody (GAH), or (3) immobilized in combination with immobilized CD3 mAb. Controls were GAH alone, or with CD3 mAb, and the CD28 IgM mAb 248.23.2. IL-2 levels were determined by
20 CTLL-2 bioassay using known amounts of IL-2 as a standard (inset).

Fig. 18 is a bar chart illustrating the level of IL-2 production by DC27.CD28wt cells treated as described in Fig. 17.

25 Fig. 19 is a bar chart illustrating IL-2 promoter activity in PCD28.1 cells stimulated as described in Fig. 17. IL-2 promoter activity was measured by induction of β -galactosidase activity which serves as a reporter gene under the control of an IL-2 promoter.

Figs. 20A and 20B are bar graphs respectively showing the induction of the IL-2 promoter, and IL-2 production by CD28 expressing cells incubated with CD80-Ig α tp, CD86-Ig α tp, or CD80-Ig.

30 Fig. 20C is a bar graph showing the levels of IL-2 production induced with soluble CD80-Ig α tp and CD86-Ig α tp in comparison to that induced by immobilized antibody to CD3.

Fig. 21 is a bar chart illustrating inhibition of biotinylated CD80-Ig α tp binding to immobilized CD28-Ig by individual compounds in the BM-34 test set. The percent inhibition range is plotted against the number of compounds showing that range of inhibition.

Fig. 22 is a profile for Superose 6 chromatography of the chimeric derivative of the
5 Epo receptor antibody 1C8 (here labeled "anti-EPOr-IgG₁") and the α tp construct of the same antibody (labeled "anti-EPOr-IgG₁ α tp") with binding activity to an immobilized EPOr-Ig protein shown in the inset.

Detailed Description of the Invention

10 The invention provides an hexameric fusion protein useful in therapeutic and immunogenic compositions. The hexameric fusion protein of the invention is particularly well suited for applications for which the binding protein from which it is derived is unsatisfactory because of low binding affinity/avidity and for other applications where
15 multivalency is desired. These applications include diagnostics, binding assays, screening assays and cellular responses based on receptor cross-linking. Also provided are compositions and methods for production and purification of these fusion proteins.

The invention further provides methods of producing stable hexameric fusion proteins, by providing a selected binding protein with an IgA tailpiece (α tp) or a functional
20 equivalent thereof. The inventors have found that addition of the α tp from the natively monomeric or dimeric IgA, surprisingly, provides the resulting fusion protein with the ability to form stable hexamers.

I. Fusion Proteins

As used herein, a hexameric fusion protein of the invention contains a dimeric
25 binding protein which has been provided at its carboxy terminus with a tailpiece (α tp) characterized by having the activity of the tailpiece from the C-terminus of the heavy chain of an IgA antibody. This tailpiece, when attached to each monomer of the dimeric binding protein, provides the resulting fusion protein with the ability to form stable hexamers, i.e., the hexameric fusion proteins of the invention do not undergo any appreciable dissociation in
30 solution (e.g., phosphate buffered saline) at room temperature.

In a particularly preferred embodiment, the fusion proteins of the invention are homo-hexamers. However, where desired, hetero-hexamers comprising two different fusion proteins may be constructed.

The binding proteins useful in the invention include full-length proteins and fragments thereof which are characterized by the binding ability of the full-length protein, i.e., the fragment which has the ability to bind to the counter-receptors or other ligands of the selected binding protein. Such binding proteins may be derived from a protein or protein
5 complex which natively dimerizes for biological activity; or may be genetically engineered as described herein. Examples of suitable natively dimeric binding proteins are those with carboxyl termini situated such that addition of the α tp to the carboxyl terminus of each polypeptide chain, with or without a linker, allows juxtaposition of the α tp chains. One of skill may readily select such native dimeric proteins or dimeric protein complexes, which
10 include, for example, IgG, IgD, or IgE antibodies, Fab fragments, Fab₂ fragments, Ig-Fc fragments, Ig fusion proteins, and the extracellular domains of cell surface proteins such as the α/β chain of a T cell receptor, CD28 and CTLA4, CD8 α/β heterodimers and α/α homodimers, and the α/β chain of integrin proteins and various cytokine receptors (e.g., IL3, IL5, etc.). These binding proteins are available from a variety of commercial and academic
15 sources. Alternatively, these sequences may be chemically synthesized.

As discussed above, a selected binding protein may be engineered to be dimeric. For example, a protein fragment comprising a binding domain of a selected monomeric binding protein may be attached to an Ig-Fc fragment which forms dimers. Desirably, the binding protein is selected from surface glycoproteins from the immunoglobulin supergene family and their ligands. For example, in a currently preferred embodiment, the binding protein is
20 selected from CTLA-4 (whose extracellular domain can be expressed as a monomer or dimer) and its counter-receptors CD80 and CD86. However, other proteins, including other binding proteins, are known to those of skill in the art and may be used in the construction of a hexameric fusion protein of the invention. Although a currently preferred embodiment of this
25 invention provides hexameric immunoglobulin fusion proteins, which are exemplified herein, this invention is not so limited. For example, a binding protein may be genetically modified to alter its activity. For example, engineered, mutant forms of IL4 have been described that retain high affinity for its receptor but lack normal agonist activity and serve as antagonists of IL-4 mediated function [see, e.g., N. Kruse et al, EMBO J., 11:3237-3244 (1992) and
30 WO96/04388 (Feb. 15, 1996)]. Such a mutant would be useful in a hexameric IL4-Ig fusion protein according to the invention, serving as an antagonist of IL4 function.

The protein fragment used to construct a dimeric binding protein contains at least a fragment of the extracellular domain of the selected binding protein. For functional binding

activity, this extracellular fragment preferably contains the sequences required for binding, which can be readily determined by one of skill in the art. In a preferred embodiment, which makes use of a eukaryotic production system, the protein fragment also contains an export leader sequence which is native to the binding protein selected. However, other export leader sequences which are capable of exporting the protein may be substituted by one of skill in the art. In one exemplary embodiment, where the target is CD28, the protein fragment is the native leader and extracellular domain from CD80 or CD86. The fragments can be obtained from proteins such as CD80 [P. S. Linsley *et al.*, J. Exp. Med., 173(3):721-730 (1991); Truneh *et al.*, Mol. Immunol., 33(3):321-334 (1996); J. E. Ellis *et al.*, J. Immunol., 56:2700-2709 (1996)], and CD86 [P. S. Linsley *et al.*, Immunity, 1:793-801 (1994); J. E. Ellis *et al.*, cited above; P. S. Linsley *et al.*, J. Exp. Med., 174:561-569 (1991)]. In another embodiment, where the target is CD80 or CD86, the protein fragment is the native leader and extracellular domain from CTLA-4 or CD28.

The Fc fragment used in the construction of the hexameric fusion protein may be from any antibody subclass, except IgA. Thus, the Fc fragment may be derived from the IgG, IgD, or IgE subclass. When the Fc fragment is derived from an IgG antibody, any of the human isotypes, i.e., IgG₁, IgG₂, IgG₃, and IgG₄, may be selected. Further, the parental IgG antibody may be mutated to reduce binding to complement or Ig-Fc receptors [see, e.g., A.R. Duncan *et al.*, Nature, 332:563-564 (1988); A. R. Duncan and G. Winter, Nature, 332:738 (1988); M.-L. Alegre *et al.*, J. Immunol., 148:3461-3468 (1992); M-H Tao *et al.*, J. Exp. Med., 178:661-667 (1993); V. Xu *et al.*, J. Biol. Chem., 269:3469-3474 (1994)]. When the Ig-Fc fragment is derived from IgM, it desirably contains the hinge/CH2/CH3/CH4 sequence, but not the naturally occurring 18 amino acid tailpiece (μ tp).

Optionally, the C-terminal end of the IgG₁ CH3 domain of the Fc fragment may be modified by conventional techniques to contain a restriction enzyme site for convenient cloning of the tailpiece segments (i.e., the peptide of the invention). Such modifications are described in more detail in the examples below, and are well known to those of skill in the art.

The peptide used to construct the fusion protein of the invention is derived from tailpiece located at the C-terminus of the heavy chain of an IgA antibody. In a preferred embodiment, this peptide is 18 residues in length and is the α tp segment of the human IgA1 heavy chain or a functional equivalent thereof. One particularly suitable peptide is: PTHVNVSVVMAEVDGTCY [SEQ ID NO: 3]. If desired, this peptide may be modified to remove the glycosylation site by changing 1 or 2 amino acids at residues 5-7 (NVS). For

example, the N (asparagine) may be changed to Q (glutamine) and/or the S (serine) may be changed to A (alanine). Additionally, up to about 4 amino acid residues of the human IgA CH3 domain may be retained. Alternatively, functional equivalents of the human IgA1 α tp may be selected. Suitable functional equivalents include, for example, gorilla IgG1, human
5 IgA2, rabbit IgA, and mouse IgA. Such functional equivalents may also be modified by removal of glycosylation sites. As described herein, this peptide is linked, directly or indirectly, to the binding protein (e.g., the Ig-Fc fragment) and provides the fusion protein of the invention with the ability to assemble into a stable hexamer.

The fusion protein may contain a linker sequence. Optionally, such a linker may be
10 located between the binding protein (e.g., the Ig-Fc fragment) and the α tp peptide. This linker is preferably an amino acid sequence between about 1 and 20 amino acid residues, and more preferably between about 1 and 12 amino acid residues, in length. Other appropriate or desired linkers may be readily selected by one of skill in the art. Although currently less desired, one of skill in the art may substitute other linkers for the preferred amino acid
15 sequence linkers described above.

Three currently preferred embodiments of the fusion proteins of the invention are described herein, CD80-Ig α tp, CD86-Ig α tp and CTLA4-Ig α tp. These proteins are composed of the native leader and extracellular domains of the CD80 (B7.1), the CD86 (B7.2, B70), and the CTLA-4 surface glycoproteins, respectively, linked to the hinge/CH2/CH3 region of
20 the heavy chain of human IgG₁ (Fc fragment) and terminating in a short tail piece segment from human IgA1 (α tp). Another example of a hexameric protein of the invention is an IgG antibody, where the α tp is joined directly to the carboxy terminus of the heavy chain and a light chain is paired with this heavy chain. The α tp hexameric antibody and Ig fusion proteins of the invention are advantageous over IgM antibodies and IgM fusion proteins in
25 that the hexamers of the invention are readily purified on commercially available chromatography supports and are more efficiently expressed.

These constructs may be made using known techniques. A detailed description of the construction of these exemplary fusion proteins of the invention is provided in the examples below.

30 Briefly, each chain of a dimeric binding protein is selected or constructed. For example, one preferred binding protein is a recombinant immunoglobulin containing the native leader and extracellular domain fused to an Ig-Fc fragment from the selected human IgG antibody. The α tp is added, optionally by introducing a convenient restriction

endonuclease site near the C-terminus of the binding protein (e.g., an Fc region) using silent mutations of the coding sequence and then cloning a synthetic oligonucleotide into this site that encodes the tailpiece segment. The tailpiece segment is matched to that of the human α -1 chain. The tailpiece provides the fusion protein with the ability to form hexamers and the
5 resulting construct is the hexameric fusion protein of the invention. A schematic representation of the predicted hexamer for an exemplary fusion construct of the invention, CD80-Ig α tp, is shown in Fig. 1.

Preferably, the fusion proteins of the invention are produced using recombinant techniques. Desirably, the nucleic acid sequences may be fused and the fusion protein
10 expressed in vitro in a suitable host cell. Alternatively, the fusion proteins of the invention are produced by separately expressing, or co-expressing the nucleic acid sequences encoding the protein fragments and α tp fragment of the invention and fusing the expressed products. Suitably, the resulting fusion protein forms hexamers. These production techniques are discussed in more detail below.

15

II. Polynucleotide Sequences, Expression and Purification

The present invention further encompasses polynucleotide sequences encoding the fusion proteins of the invention. In addition to the DNA coding strand, the nucleic acid sequences of the invention include the DNA (including complementary DNA) sequence
20 representing the non-coding strand and the messenger RNA sequence. Variants of these nucleic acids of the invention include variations due to the degeneracy of the genetic code and are encompassed by this invention. Such variants may be readily identified and/or constructed by one of skill in the art. Further, the polynucleotide sequences may be modified by adding readily assayable tags to facilitate quantitation, where desirable.

25 To produce recombinant fusion proteins of this invention, the DNA sequences of the invention are inserted into a suitable expression system, preferably a eukaryotic system. Desirably, a recombinant vector is constructed in which the polynucleotide sequence encoding at least one chain of the fusion protein (i.e., the binding protein/ α tp) is operably linked to a heterologous expression control sequence permitting expression of the fusion protein of the
30 invention. Numerous types of appropriate expression vectors and host cell systems are known in the art for expression, including, e.g., mammalian, yeast, bacterial, fungal, drosophila, and baculovirus expression.

The transformation of one or more of these vectors into appropriate host cells results in expression of the fusion proteins of the invention. Other appropriate expression vectors, of which numerous types are known in the art, can also be used for this purpose.

Such production methods permit assembly of the hexameric fusion protein of the invention by the host cell. Typically, such methods will provide a homo-hexameric fusion protein. However, in another embodiment, hexameric fusion proteins of mixed specificity may be produced by co-expression of different fusion proteins (i.e., binding protein/ α tp). For example, two fusion proteins recognizing non-competing sites on the same molecule can be co-expressed resulting in hexamers that can bind to two sites on the same molecule, resulting in higher binding avidity than for each fusion protein alone or as a homogenous hexamer. Alternatively, the two fusion proteins can bind to two distinct molecules presented on the same, or different surfaces (e.g., expressed on the same or different cells).

Suitable host cells or cell lines for transfection by this method include mammalian cells, such as Human 293 cells, Chinese hamster ovary cells (CHO), the monkey COS-1 cell line, murine L cells or murine 3T3 cells derived from Swiss, Balb-c or NIH mice. Suitable mammalian host cells and methods for transformation, culture, amplification, screening, and product production and purification are known in the art. [See, e.g., Gething and Sambrook, Nature, 293:620-625 (1981), or alternatively, Kaufman *et al.*, Mol. Cell. Biol., 5(7):1750-1759 (1985) or Howley *et al.*, U. S. Patent 4,419,446]. Another suitable mammalian cell line is the CV-1 cell line.

Other host cells include insect cells, such as *Spodoptera frugiperda* (Sf9) cells. Methods for the construction and transformation of such host cells are well-known. [See, e.g. Miller *et al.*, Genetic Engineering, 8:277-298 (Plenum Press 1986) and references cited therein].

Although less preferred, also useful as host cells for the vectors of the present invention are bacterial cells. For example, the various strains of *E. coli* (e.g., HB101, MC1061) are well-known as host cells in the field of biotechnology. Various strains of *B. subtilis*, *Pseudomonas*, other bacilli and the like may also be employed in this method.

Many strains of yeast cells known to those skilled in the art are also available as host cells for expression of the proteins of the present invention. Other fungal cells may also be employed as expression systems.

Thus, the present invention provides a method for producing a fusion protein of the invention which involves transforming a host cell, preferably a eukaryote, with at least one expression vector containing a recombinant polynucleotide encoding a fusion protein under the control of a transcriptional regulatory sequence, e.g., by conventional means such as
5 transfection or electroporation. The transformed host cell is then cultured under suitable conditions that allow expression of the fusion protein. The expressed and assembled fusion protein is then recovered, isolated, and purified from the culture medium by appropriate means known to one of skill in the art. In a preferred embodiment, the fusion proteins are assembled by the host cell following co-production of one or more of the fusion proteins of
10 the invention. Alternatively, the hexameric fusion protein may be assembled following recovery from the host cell.

Advantageously, the fusion proteins of the invention can be readily purified using conventional techniques. For example, hexameric Ig fusion proteins of the invention may be readily purified on high affinity, high capacity supports based on protein A and protein G.
15 Such resins are commercially available [Pharmacia Inc.; Bioprocessing Ltd.].

Although less preferred, the hexameric fusion protein may be produced in insoluble form. For example, the proteins may be isolated following cell lysis in soluble form, or extracted in guanidine chloride.

20 III. Pharmaceutical Compositions and Methods of Use Thereof

The fusion proteins of this invention or DNA sequences encoding them may be formulated into pharmaceutical compositions and administered using a therapeutic or immunogenic regimen compatible with the particular formulation. Pharmaceutical compositions within the scope of the present invention include compositions containing a
25 protein of the invention in an effective amount to have the desired physiological effect.

Suitable formulations for parenteral administration include aqueous solutions of the active compounds in water-soluble or water-dispersible form, e.g., saline. Alternatively, suspensions of the active compounds may be administered in suitable conventional lipophilic carriers or in liposomes. In still another alternative, adjuvants may be desired, particularly
30 where the composition is to be used as an immunogen.

The compositions may be supplemented by active pharmaceutical ingredients, where desired. Optional antibacterial, antiseptic, and antioxidant agents in the compositions can perform their ordinary functions. The pharmaceutical compositions of the invention may

further contain any of a number of suitable viscosity enhancers, stabilizers, excipients and auxiliaries which facilitate processing of the active compounds into preparations that can be used pharmaceutically. Preferably, these preparations, as well as those preparations discussed below, are designed for parenteral administration. However, compositions designed
5 for oral or rectal administration are also considered to fall within the scope of the present invention.

As used herein, the terms "suitable amount" or "effective amount" means an amount which is effective to treat or prevent the conditions referred to below. A preferred dose of a pharmaceutical composition containing a fusion protein of this invention is generally effective
10 above about 0.1 mg fusion protein of the invention per kg of body weight (mg/kg), and preferably from about 1 mg/kg to about 100 mg/kg. These doses may be administered with a frequency necessary to achieve and maintain satisfactory fusion protein levels. Although a preferred range has been described above, determination of the effective amounts for treatment or prophylaxis of a particular condition may be determined by those of skill in the
15 art.

Particularly, pharmaceutical compositions containing the hexameric antibody/ α tp fusion proteins of the invention are useful as antagonists for the 7 transmembrane (7 TMR) class of cell surface receptors, since such receptors are often arrayed in many copies on cell surfaces and the aggregation of such receptors does not lead to intracellular signalling (agonism) as can occur for many other types of cell surface receptors. For example,
20 administration of a pharmaceutical compositions containing a hexameric antibody/ α tp fusion protein of the invention blockades chemokine receptors, a subfamily of the 7 TMR, and inhibits chemotaxis and activation of target cells such as eosinophils. A second example is CTLA4-Ig α tp. CTLA4-Ig is a potent inhibitor of CD80 and CD86 driven stimulation of
25 T-cells through their interaction with CD28. In animal models, CTLA4-Ig has shown benefit in several autoimmune diseases and transplantation.

Since the CD80 and CD86 antigens recognized by CTLA4-Ig are arrayed in many copies on the cell surface, an α tp hexameric form of CTLA4-Ig may provide a more potent antagonist than the standard Ig fusion protein. In another embodiment, a pharmaceutical
30 composition of the invention containing Ig α tp fusion proteins of the invention may be used for removal of complement components or components of the blood coagulation cascade to retard clotting.

In one aspect, the invention provides a method for antagonizing cell surface CD80- and CD86-mediated stimulation of CD28 positive cells by administering to the cells a hexameric fusion protein CTLA4-Ig α tp. This may be performed *in vivo*, by administering a pharmaceutical composition containing this hexameric fusion protein. In another aspect, the invention provides a method for stimulating (agonist activity) CD28+ T cells by administering the CD80- or CD86-hexameric fusion protein to the cells in culture resulting in stimulation of IL-2 production from these cells. These proteins may be used alone, or in combination with other stimulators of T-cells (e.g., antibodies directed against the T cell receptor-CD3 complex.)

10 In another embodiment, the compositions of the invention containing Ig-Fc-containing fusion proteins are useful for *in vivo* clearance of soluble ligands, in view of the fact that hexamerization of the Fc domain enhances interaction with complement components and Fc receptors. Thus, ligands bound to the hexameric fusion protein of the invention are efficiently cleared from circulation.

15 The hexameric fusion proteins of the invention can also serve as agonists, particularly in situations where aggregation can induce a desired response. For example, aggregation is essential for signal transduction through many cell surface receptors - either as a consequence of multivalent presentation of the receptor ligand (eg., a counter receptor on a the surface of a second cell) or through changes induced upon ligand binding, or both. An example of signalling through a cell surface receptor induced by cross-linking through recognition of its counter-receptor on a second cell is CD28 recognition by CD80 or CD86.

Thus, the invention further provides a method for stimulating CD28 positive cells by administering to CD28 positive cells CD80-Ig α tp and/or CD86-Ig α tp. Examples of soluble ligands inducing signal transduction through binding to their receptors are EGF and growth hormone and both result in receptor dimerization. For these receptors, dimerization induced through antibody binding also can lead to activation [Schreiber *et al.*, Proc. Natl. Acad. Sci. USA, 78:7535 (1981), Fuh *et al.*, Science, 256:1677 (1992)]. Hexameric antibodies against such receptors or hexameric ligand-Ig fusion proteins for these receptors are expected to be more efficient stimulators than the standard dimeric antibodies or ligand Ig fusion proteins. For example, the pharmaceutical compositions containing the hexameric antibodies or cytokine-Ig fusion proteins of the invention are useful in inducing signal transduction in receptors for hematopoietic cytokines, such as erythropoietin, thymopoietin and growth stimulatory factor.

Also provided is a method for suppressing CTLA-4 positive cells by administering CD80-Ig α tp and/or CD86-Ig α tp to CTLA4 positive cells. This may be performed *in vivo*, by administration of a pharmaceutical composition containing the hexameric proteins. Alternatively, the hexameric proteins are added to CTLA4 positive T-cells in culture resulting in inhibition of IL-2 production from these cells.

In yet another aspect, hexameric Ig-fusion proteins of the invention can also serve as enhanced immunogens for the fused protein fragment due to efficient, receptor-mediated uptake for antigen processing and presentation or efficient interaction with proteins of the complement system. Enhanced immunogenicity is desirable for the efficient generation of polyclonal and monoclonal antibodies and for therapeutic vaccination. Thus, the invention further provides a method of immunizing using the pharmaceutical composition of the invention.

IV. Assays

The hexameric fusion proteins of the invention are useful in *in vitro* assays for measuring the binding of the fusion protein to a selected ligand and for identifying the native or synthetic ligand for the binding proteins. Such a ligand includes the native ligand or counter-receptor to the binding protein from which the hexameric fusion protein is derived. For example, where the fusion protein is derived from CD80 or CD86, the ligand may be CD28 or CTLA-4. Alternatively, the ligand may be a derivative of the native counter-receptor, a peptide, peptide-like compound, or a chemical compound which interacts with the fusion protein.

The hexameric fusion proteins may be used for *in vivo* assays, including, for example imaging. See, e.g., S. M. Larson *et al.*, *Acta Oncologica*, 32(7-8):709-715 (1993); R. DeJager *et al.*, *Seminars in Nuclear Medicine*, 23(2):165-179 (Apr. 1993).

Alternatively, a fusion protein of the invention may be used to screen for new ligands. The use of the fusion proteins of this invention in such an assay is particularly well suited for identifying cell surface or multivalent ligands.

Suitable assay methods may be readily determined by one of skill in the art. For example, an ELISA format may be utilized in which the selected ligand is immobilized, directly or indirectly (e.g., via an anti-ligand antibody) to a suitable surface.

Where desired, and depending on the assay selected, the hexameric fusion protein may be immobilized on a suitable surface. Such immobilization surfaces are well known.

For example, a wettable inert bead may be used in order to facilitate multivalent interaction with the hexameric fusion proteins of the invention.

Further, the methods of the invention are readily adaptable to combinatorial technology, where multiple molecules are contained on an immobilized support system.

5 Thus, the fusion proteins of the invention permit screening of chemical compound and peptide based libraries where these agents are presented in a multivalent format compatible with more than one subunit of the hexamer. Monomeric interactions of this type are routinely in the mM range and thus may not be readily detected with monomeric proteins. Advantageously, the avidity of the hexameric fusion proteins of the invention permit direct binding.

10 Typically, the surface containing the immobilized ligand is permitted to come into contact with a solution containing the fusion protein and binding is measured using an appropriate detection system. Suitable detection systems include the streptavidin horse-radish peroxidase conjugate, direct conjugation by a tag, e.g., fluorescein. Other systems are well known to those of skill in the art. This invention is not limited by the detection system used.

15 The assay methods described herein are also useful in screening for inhibition of the interaction between a hexameric fusion protein of the invention (and thus, the binding protein from which it is derived) and its ligand(s). For example, one may screen for inhibitors of CD80 and CD86 binding to CD28 and CTLA-4. In a preferred method, a solution containing the suspected inhibitors is contacted with an immobilized recombinant CD28 or CTLA-4
20 protein substantially simultaneously with contacting the immobilized ligand with the solution containing the hexameric CD80- or CD86-Ig α p protein. The solution containing the inhibitors may be obtained from any appropriate source, including, for example, extracts of supernatants from culture of bioorganisms, extracts from organisms collected from natural sources, chemical compounds, and mixtures thereof. In another variation, the inhibitor
25 solution may be added prior to or after addition of the CD80- or CD86-Ig α p proteins to the immobilized CD28 or CTLA-4 protein. Similar methods may be performed using other hexameric fusion proteins of the invention and their respective ligands.

The large size of the Ig α p fusion proteins is also advantageous for biophysical assay methods dependent on diffusion or rotation of the protein target in solution, such as for
30 example, fluorescence polarization, fluorescence correlation spectroscopy and anisotropic analytical methods.

These examples illustrate the preferred methods for preparing and using the fusion proteins of the invention. These examples are illustrative only and do not limit the scope of the invention.

5 Example 1 - Production and characterization of exemplary α tp Ig fusion proteins

The following describes the production of CD80-Ig α tp, CD86-Ig α tp, and CTLA4-Ig α tp. Further, for comparison, a construct containing the human IgM tailpiece added to the C-terminus of CD80-Ig was also prepared. This construct, designated CD80-Ig μ tp, differs in amino acid sequence from the α tp derivative as follows:

	<u>CH3</u>	<u>Tailpiece</u>	<u>SEQ ID NO:</u>
10 IgG1	SLSPGK	(none)	9
μ tp	SLSTGK	PTLYNVSLVMSDTAGTCY	25 and 10
α tp	SLSAGK	PTHVNVSVVMAEVDGTCY	26 and 11

15 A. Construction of Recombinant Ig: Binding Protein Fragment/Fc Fusions

The pHbactCd28neo vector for expression of CD28 was previously described [D. Couez *et al.*, Molecul. Immunol., 31(1):47-57 (1994)]. For expression of CD80, the coding sequence was cloned by PCR and inserted into a derivative [Dr. F. Letourneur, NIH] of pCDLSR α 296 [Y. Takebe *et al.*, Molecul. & Cell. Biol., 8(1):466-472 (1988)] as described
 20 [C. A. Fargeas *et al.*, J. Exp. Med., 182:667-675 (1995)].

The vector COSFcLink [A. Truneh *et al.*, Mol. Immunol., 33(3):321-334 (1996)] was constructed for expression of proteins C-terminally fused to a human IgG1 Fc region under the transcriptional control of the major late promoter of CMV. The dhfr cassette in this vector permits selection for gene amplification in response to methotrexate. The coding
 25 sequences for the native leader and extracellular domain peptide of CD28 and CD80 were grafted onto a human IgG1 heavy chain Fc region in the vector COSFcLink, beginning at the start of the hinge region, in a manner similar to that previously described for CD28 and CD80 [P.S. Linsley *et al.*, J. Exp. Med., 174(3):561-569 (1991)]. The Fc region in this vector was derived from the human plasma leukemia cell line ARH-77 [ATCC CRL 1621]
 30 and contains a mutation of cysteine to alanine in the upper hinge region (SEQ ID NO: 27 EPKSA, where the mutation is underscored). The CD28 and CD80 sequences were cloned as KpnI - Eag I fragments by PCR from the vectors described above and inserted into the corresponding sites in COSFcLink. The resulting vectors are termed CD28FcLink and

CD80FcLink, respectively. For CD28-Ig, the junction of receptor/Fc fragment (immunoglobulin junction) is SEQ ID NO: 12 --GPSKP/EPKSA-- and the mature processed N-terminal sequence is SEQ ID NO: 13 NKIL --. For CD80-Ig, the immunoglobulin junction is SEQ ID NO: 14 -HFPDq/EPKSA-- and the mature processed N-terminal sequence is VIHV-- (Fig. 4A-4D). The lower case "q" in CD80 represents the substitution of glutamine for the native asparagine.

CD86-Ig, the corresponding binding protein/Fc construct for CD86 containing the native signal peptide of CD86 (B70) [M. Azuma *et al.*, Nature, 366:76-79 (1993)], was constructed using methods essentially identical to those described above. The signal and extracellular sequences were PCR cloned from a plasmid containing the CD86 (B70) coding region that was obtained by reverse transcriptase/PCR cloning from human B-cell RNA based on the sequence described by M. Azuma *et al.* (above). Sequence analysis confirmed identity of this cloned CD86 (B70) region with that of Azuma *et al.* (above). The amino acid sequence at the junction to the Fc region is: SEQ ID NO: 16 --PPPDHepksa-- where capital and lower case letters indicate CD86 and Fc sequences respectively. The mature processed N-terminal sequence is SEQ ID NO: 17 LKIQ -- (Fig. 5A-5B).

CTLA4-Ig, the corresponding binding protein/Fc construct for human CTLA4 containing the native signal peptide of CTLA4 [P. Dariavach *et al.*, Eur J Immunol, 18: 1901-1905 (1988); Harper *et al.*, J Immunol, 147: 1037-1044 (1991)] was constructed in a similar manner. HuC4.32, a pCDM8 plasmid containing the cDNA sequence for human CTLA4 (Harper *et al.*, above) was provided by the laboratory of P. Golstein (Centre d'Immunologie INSERM-CNRS de Marseille-Luminy, 13288 Marseille Cedex 9, France). For PCR cloning of the extracellular domain, the 5' primer was positioned in the pCDM8 vector. [Abberent cloning led to deletion of about 140 bp upstream of the EcoRI site relative to CD80FcLink (compare Fig. 6 with Fig. 3, below).] The amino acid sequence at the junction spanning the end of the CTLA4 extracellular domain and the hinge region is: SEQ ID NO: 18 --EPCPDSDAepksa-- where capital and lower case letters indicate CTLA4 and Fc sequences respectively and the underlined alanine residue indicates its substitution for phenylalanine in the native CTLA4 sequence. The mature processed N-terminal sequence is SEQ ID NO: 19 MHVA-- (Fig. 6A-6C).

Hexameric forms of the CTLA4, CD80 and CD86 recombinant Ig proteins were created by addition of a sequence encoding the 18 amino acid tail piece region of human IgA1

heavy chain to the C-terminus of the CH3 domain in the expression vectors described above. These methods are described in detail below.

B. Construction of Hexameric Fusion Proteins

For convenience, a Hind III site was introduced into the CH3 domain of
 5 CD80FcLink [spanning the 3rd base of the codon for Leu441 [EU numbering, E.A. Kabat *et al.*, cited above] through the 2nd base of the codon for L443]. The Hind III site was introduced by standard PCR methods (eg., PCR Protocols: A Guide to Methods and Applications, Innis *et al.*, eds, 1990) using the following oligonucleotides:
 5' oligo (positioned in the hinge region of the vector): SEQ ID NO: 20

10

EagI

cccaaatcggccgacaaaact

3' oligo (spanning the C-terminus of CH3): SEQ ID NO: 21

15

XbaI

HindIII

tcagcgagctctagactacactcatttaccggagacaagcttaggctcttctgcgt

The PCR fragments were isolated by agarose gel electrophoresis and purified on Spin Bind columns (FMC Corp). The fragment was digested with Eag I and Xba I and cloned into
 20 similarly digested CD80FcLink vector and colonies were screened for the newly created Hind III site, yielding the vector CD80FcLink-Hd.

To introduce the α tp sequence, a synthetic oligonucleotide linker encoding this sequence was cloned between the newly created Hind III site and the Xba I site in CD80FcLink-Hd. The complementary oligonucleotides for the linker sequence were:

25 (5') SEQ ID NO: 22 agcttgctgcgggtaaacccacccatgtcaatgtgtctgttgcaggc
 (Hind III adaptor)

ggaggtggacggcacctgctactgatagt

(5') SEQ ID NO: 23 ctagactatcagtagcaggtgccgtccacctccgcatgacaacagac
 30 (Xba I adaptor)

acattgacatgggtgggtttaccgcagaca

5 mg of each linker was denatured at 70°C for 10 minutes. The reactions were cooled to room temperature for 20 minutes. The concentration of linker was titrated from 50 to 5 ng using 1000 ng of gel purified CD80FcLink-Hd vector, digested with Hind III/Xba I. Several colonies from each ligation condition were screened for the presence of the α p linker by PCR and confirmed by DNA sequencing.

A schematic representation of the resulting vector, CD80Fc α pLink, is shown in Fig. 2 and the complete DNA sequence is given in Fig. 3A-3H. The vector sequence may differ in some sites from the actual plasmid, but would be function. Introduction of the CD80-Ig α p coding region into other standard mammalian expression vectors (e.g., pBK-CMV from Stratagene, La Jolla, CA) will give suitable results and can be modified appropriately, e.g., by introduction of a dhfr gene, by one of skill in the art.

The vectors for expression of CD86-Ig α p as derived from the corresponding Ig expression vector by replacing the Fc coding region with the Fc- α p region from CD80-Ig α p. The Fc segment of CD86FcLink was excised by cleavage with Eag I (in the hinge region) and Xba I (following the C-terminus of CH3) and replaced with the corresponding fragment of CD80Fc α pLink to give the expression vector CD86Fc α pLink. The vector for expression of CTLA4-Ig α p was derived by replacing a SpeI-EagI fragment in CD80Fc α pLink with the corresponding fragment from CTLA4FcLink to give the expression vector CTLA4Fc α pLink. The SpeI site is at base 46 in the CMV promoter region. The sequences of the CD86 and CTLA4 constructs in the region differing from CD80Fc α pLink are given in Figs. 5 and 6.

By a similar approach a vector encoding CD28-Ig α p could be prepared starting the CD28FcLink vector described in part A above, or a similar construct encoding an altered version of the CD28 extracellular sequence.

C. Production and purification

The CD28-Ig, CD80-Ig and CD86-Ig proteins were produced in CHO cells and purified as described in A. Truneh et al., *Mol Immunol*, 33: 321-334 (1996) and in I. Kariv et al., *J Immunol*, 157: 29-38 (1996). The CTLA4-Ig protein was produced and purified in a similar manner, using the vector construct described above in part A of this section. The Ig α p fusion proteins were shown to be produced upon transfection of the Fc α pLink vectors into COS-7 cells following standard procedures for transfection of COS cells (eg., Current Protocols in Molecular Immunology, edited by F.M. Ausubel *et al.* 1988,

John Wiley & Sons, vol 1, section 9.1) and for immunoblot analysis (eg., JR Jackson *et al.*, J. Immunology, 154:3310-3319 (1995)) with rabbit polyclonal anti-sera prepared against various derivatives of the CD80, CD86, and CTLA4 proteins or goat anti-human Fc antibody. The α tp and μ tp constructs of CD80-Ig were compared in terms of their efficiency of expression and oligomerization. As determined by SDS/PAGE and immunoblot analysis, the CD80-Ig μ tp construct did not express as well as the α tp construct of the invention (not shown). The α tp and μ tp proteins were purified from the COS cell supernatants by capture on Prosep A (Bioprocessing, Ltd., Consett County Durham, U.K.) and their state of oligomerization examined by analytical size exclusion chromatography on a 3.2 X 30 mm Superose 6 column run on a Smart System HPLC (Pharmacia Biotech, Piscataway NJ). Both proteins showed a similar profile of a dominant large MW species eluting in the molecular weight range of IgM, consistent with formation of a hexameric structure, and a smaller fraction that eluted at the same size as CD80-Ig itself (not shown). However, the fraction of apparent hexamer in the α tp construct was higher (about 80%) than for the μ tp construct (about 60%). Both the higher level of expression and the greater efficiency of oligomer formation indicated that the α tp construct of the invention was superior to the μ tp derivative. Subsequently, the CD86-Ig α tp and the CTLA4-Ig α tp proteins were produced in COS cells at about the same level observed for the CD80-Ig α tp protein (0.1 - 0.2 ug/ml). The CD80- and CD86-Ig α tp proteins were then produced in a CHO cell system (A. Truneh *et al.*, Mol Immunol, 33: 321-334 (1996)) at levels of 5-10 mg/L. This level of production is comparable to other highly expressed proteins (e.g. antibodies) produced in the same manner in this system.

These results indicate that development of standard amplified CHO cell lines with high production levels of hexamer (50 mg/L or greater) is feasible. A procedure for transfection and amplification in CHO cells is described in P. Hensley *et al.*, J. Biol. Chem., 269:23949-23958 (1994)). Briefly, a total of 30 ug of linearized plasmid DNA (e.g. CD80Fc α tp μ link) is electroporated into 1×10^7 cells. The cells are initially selected in nucleoside-free medium in 96 well plates. After three to four weeks, media from growth positive wells is screened for expression - e.g., in an ELISA format using an antibody directed against the Fc region of human IgG1. The highest expressing colonies are expanded and selected in increasing concentrations of methotrexate for amplification of the transfected vectors. If a commercial vector like pBK-CMV (noted above) is used, a dhfr gene should be

introduced into this plasmid or provided on a second co-transfecting plasmid to allow selection of amplification in methotrexate.

The proteins produced in CHO cells were purified by protein A affinity and size exclusion chromatography. For the CD80-Ig hexamer, thirty liters of conditioned medium containing CD80-Ig α tp were chromatographed on a Protein A Sepharose Fast Flow column (Pharmacia) at 20 ml/min. The column (5.0 x 11.6 cm; 225 ml) were preequilibrated in 20 mM sodium phosphate, 150 mM NaCl, pH 7.5 (PBS). After loading, the column was washed with 1.8 L of PBS to baseline absorbance. CD80-Ig α tp was eluted with 0.1 M sodium citrate, pH 3.0 at 10 ml/min. The eluate was neutralized immediately with 1 M Tris-HCl, pH 8.0. After filtration with a Sterivex GV filter (Millipore) using a 60 ml syringe, CD80-Ig α tp was concentrated using an Amicon stirred cell and a YM100 membrane to 1.3 mg/ml. CD80-Ig α tp was frozen using a dry ice ethanol bath and stored at -70°C.

To separate hexamer from monomer, 10 ml of the concentrated CD80-Ig α tp was chromatographed on a Superdex 200 column (2.6 x 60 cm; Pharmacia) at 2.5 ml/min. The first peak (eluted at about 45 minutes) containing the majority (about 90%) of the 280 nm absorbing material was pooled (20 ml; 0.6 mg/ml), frozen as before and stored at -70 (Figure 7). This material eluted at approximately the position of thyroglobulin (~700,000 Da.) just behind the void volume. A minor peak at about 57 minutes corresponded to "monomer" CD80-Ig. The integrity of the CD80-Ig α tp in the peak fractions is shown by the single band observed in coomassie stained SDS/PAGE gel run under reducing conditions (lane R in the inset in Figure 7). The diffuse nature of the band is characteristic of highly glycosylated proteins and is thus expected for CD80-Ig α tp which contains 10 consensus N-linked glycosylation sites per polypeptide chain. Under nonreducing conditions, all of the protein migrates as high molecular weight species (lane NR in Figure 7, insert). This dominant fraction migrated as a symmetrical peak at a MW consistent with a hexamer with a lesser amount of a species that migrated at the size observed for the monomeric CD80-Ig protein (i.e., the Ig homodimer). N-terminal amino acid sequence analysis revealed identity to the previous analysis of CD80-Ig and to that described by others [G. J. Freeman *et al.*, 174(3):625-631 (1991)]. The CD86-Ig α tp protein was purified in a similar manner. The CTLA4-Ig α tp protein was expressed in COS cells, but not further characterized.

D. Protein characterization - Molecular Size

The size exclusion chromatography noted above during purification was consistent with formation of a homogeneous hexameric species containing six CD80-Ig subunits. The size and homogeneity of the CD80-Ig α tp protein produced in CHO cells was also investigated by analytical ultracentrifugation. Equilibrium sedimentation data for CD80-Ig α tp in PBS, pH 7.4 is shown in Fig. 8, lower panel. The sample was sedimented at 6000 rpm for 87 hours at 25 °C in a Beckman XL-A analytical ultracentrifuge. The weight average molecular weight for a fit to all the data was 1,125,000 +/- 5,000 Da. The expected molecular mass of the hexamer of 864,000, assuming 2000 Da. for each N-linked glycosylation site. The data could also be fitted to a hexamer \leftrightarrow (hexamer)₂ model with a K_d of $\sim 2 \times 10^{-7}$ M. The curves in the lower panel are for the fitted distribution of hexamer and (hexamer)₂. The sum of these two curves fits the observed data well. Inclusion of terms for a monomer (131 kDa) did not improve the fit. The distribution of residuals (fitted-observed data) for the fit of the monomer dimer model to the data is shown in the upper panel of Fig 8. The residuals are small and random, indicating a good fit. For a description of the analysis see W. Chan *et al.*, Folding and Design, 1(2): 77-89 (1996). The lower panel inset shows $g(s^*)$ analysis of velocity sedimentation data of the protein taken in the absorption mode. Data was collected at 30,000 rpm at 22 °C. The data could be fitted to two species, one of 19.4 S and one of 26.7 S which could be the hexamer and (hexamer)₂ species. For $g(s^*)$ data analysis, see W. F. Stafford, Current Opinion in Biotechnology, 8(1): 14-24 (1997).

The size and extent of covalent association of the CD80- and CD86-Ig α tp proteins were examined by SDS/PAGE. Under reducing conditions all of the protein migrated in a diffuse band at about the same size as the corresponding standard Ig constructs, as shown for the CD80-Ig α tp protein in the inset in Figure 7. Under nonreducing conditions in a 4% gel, the Ig α tp constructs migrated as very diffuse bands in the size range of IgM with little material co-migrating with the corresponding Ig constructs at about 150,000 Da (not shown). These results indicate that most of the individual polypeptide chains in the Ig α tp proteins are covalently joined through cystine bonds, consistent with the described disulfide bond formation among the cysteine residues in the μ tailpiece segment of IgM [A. C. Davis *et al.*, EMBO J., 8(9): 2519-2526 (1989)]. The diffuse nature of the high molecular weight

bands may reflect incomplete disulfide bond formation but also is expected since a hexamer form of CD80- or CD86-Ig α tp would contain 120 potential N-linked glycosylation sites.

E. Protein Characterization - Binding properties

In several assay formats the hexameric CD80- and CD86-Ig α tp proteins
5 were distinguished from the corresponding standard Ig fusion proteins by their markedly higher binding avidity to CD28 when it was presented in a multivalent array.

1) Binding to immobilized CD28-Ig in an ELISA format

For this assay format, the CD80-Ig α tp protein was biotinylated for simplicity of assay and for ease of detection since the CD28 protein absorbed to the plate wells was also a human Ig
10 fusion construct. Biotinylation was carried out essentially as described in Avidin-Biotin Chemistry: A handbook, M. D. Savage *et al.*, Pierce Chemical Company (1992). In several preparations of the protein, the molar ratio of biotin/CD80-Ig monomer was about 10:1. All steps of the assay after coating were carried out at room temperature.

The wells of 96 well microtiter plates (Immunlon 4, Dynatech Laboratories) were
15 coated with CD28-Ig (1, 2, or 4 μ g/ml) in 100 μ l/well of 0.1 M sodium bicarbonate, pH 9.4 and incubated overnight @ 4°C. The wells were washed with PBS (phosphate buffered saline) and blocked with 0.5% gelatin in PBS for 1 hour. Following an additional PBS wash, biotinylated CD80-Ig α tp was serially diluted in PBS containing 1 mg/ml BSA, 0.05% Tween directly in the wells in a final volume of 0.1 ml and incubated for 1 hour. The wells were
20 washed with PBS and bound CD80-Ig α tp protein was measured by the addition of 0.1 ml of streptavidin-HRP (streptavidin conjugated with horseradish peroxidase (Southern Biotech)) at a 1:2000 dilution for 1 hour, followed by washing and color development with 100 μ l ABTS substrate (Kierkegaard and Perry Laboratories Inc., Maryland) and measurement of absorbance at 405 nm. In some cases the color reactions were arrested by addition of 100 μ l
25 of 1% SDS prior to measurement of absorbance. A plot of CD80-Ig α tp binding versus concentration of added protein is shown in Fig. 9. In this figure, "CD28-Fc", "CTLA4-Fc", and "B7-FcA" denote CD28-Ig, CTLA4-Ig, and CD80-Ig α tp, respectively. These curves (Fig. 9) indicate that concentration dependent binding of biotinylated CD80-Ig α tp was inhibited by simultaneous addition of the CD28.1 MAbs (a murine MAbs to human CD28 that
30 inhibits binding of CD80 to CD28; Nunes *et al.*, Int. Immunol., 5:311-315 (1993)) or CTLA4-Ig protein (here labeled as CTLA4-Fc). Under the same conditions, biotinylated CD80-Ig itself showed little binding and only at much higher concentrations (Figure 10). In the same format biotinylated CD86-Ig α tp also showed good binding to CD28-Ig (Fig

10). All three biotinylated proteins showed good binding to immobilized CTLA4-Ig (Fig. 11), as expected because of the higher affinity of this interaction [P. S. Linsley *et al.*, Immunity 1: 793-801 (1994), and see part 4 of this example below], and the rank order of binding was the same as observed with immobilized CD28-Ig.

5 The specificity of the binding reaction was demonstrated by the expected hierarchical competition of binding with (1) CTLA4-Ig, (2) CD28.2 [Nunes *et al.*, 1993, cited above], a murine MAb to human CD28 that inhibits binding of CD80 to CD28, (3) unlabeled CD80- and CD86-Ig α tp proteins, (4) and the expected much weaker inhibition by the monomeric CD80-Ig fusion protein. One example is shown in Fig. 12. Briefly, microtiter wells were
10 coated with 2 μ g/ml CD28-Ig and biotinylated CD80-Ig α tp was added at a concentration of 50 μ g/ml, followed immediately by the indicated amounts of unlabeled CD80-Ig α tp (B7FcA), CD80-Ig (B7Ig), CTLA4-Ig, or the MAb CD28.2. At 50 μ g/ml, the biotinylated CD80-Ig α tp gives about 50% saturation of OD₄₀₅ (see Fig. 9). CD80-Ig was much less efficient than CD80-Ig α tp in blocking binding, consistent with the expected lower affinity/avidity of
15 the CD80-Ig protein for the immobilized CD28-Ig protein. The controls gave the expected results - the CD28.2 MAb blocked the binding site on CD28 and similarly, CTLA4-Ig blocked the binding sites on CD80-Ig α tp.

Other assay formats are possible. A second example utilizes a CD28-*mu*Ig fusion protein constructed in a manner analogous to CD28-Ig except that the Ig region was derived
20 from mouse Ig2a instead of human IgG1. More particularly, the protein was expressed using the vector CosCD28mFc2aLink, which is comparable to the CosCD28FcLink vector (described above), except that the human IgG1-Fc region was replaced with that of mouse IgG2a beginning at the Eag I site in the hinge sequence [described in I. Kariv *et al.*, J. Immunol., 157:29-38 (1996)]. The amino acid sequence in the resulting hybrid hinge region
25 is as follows: SEQ ID NO: 24 --GPSKPe~~ps~~ag**IKP**--, where capital letters correspond to the end of CD28 sequence, lower case letters are residues from the human IgG₁ hinge region, underlined lower case letters are a 2 residue substitution introduced to create an Eag I site, and bold capital letters indicate the beginning of murine IgG2a hinge region.

The CD28-*mu*Ig protein was indirectly immobilized in wells using goat anti-mouse
30 Fc antibody and then CD80-Ig α tp binding was carried out similarly to that described above. More specifically, CD28-*mu*Ig proteins containing wild-type or mutant CD28 sequences and, at equal concentrations, were captured on goat anti-mouse IgG antibody coated 96 well plates. The plates were washed with 1x PBS, blocked with 0.5% gelatin-PBS for 1 hour, and

then incubated with either biotinylated CD80- or CD86-Ig α tp for 45 min. The plates were washed and Ig α tp fusion protein was quantitated as described above. This assay was used to examine the effects of mutations in CD28 on binding to CD80 and CD86, as illustrated in Figs. 13A and 13B (I. Kariv *et al.*, J. Immunol., 157:29-38 (1996)). Each of the mutant

5 CD28-muIg2a proteins was captured on the goat anti-mouse IgG coated wells and the binding of biotinylated CD80 - Ig α tp (Fig. 13A) or CD86 - Ig α tp (Fig. 13B) was measured. Equivalent capture of each of the CD28-muIg2a proteins was verified by the comparable binding of polyclonal rabbit CD28 antisera to each of the proteins (Fig. 13C).

10 2) Binding to immobilized CD28-Ig in a biosensor assay format

The binding of CD80-Ig α tp or CD80-Ig to immobilized CD28 were compared by surface plasmon resonance analysis using a BIAcore instrument, following procedures similar to that described for other proteins [K. Johanson *et. al.*, J. Biol. Chem., 270: 9459-9471 (1995), and references therein].

15 For comparison of CD80-Ig and CD80-Ig α tp, approximately 4000 RU of CD28-Ig were immobilized onto a BIAcore CM5 sensor surface (BIAcore, Piscataway, NJ) by covalent attachment to the surface through its amines. Covalent attachment was achieved by firstly activating the surface with a 1:1 mixture of 0.1 M solution of N-hydroxysuccinamide and 0.1 M 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide. A

20 solution of CD28-Ig 50 ug/ml in 0.01 M sodium acetate pH 4.7 was then passed over the surface. Unreacted N-hydroxysuccinamide esters were then deactivated with 1 M ethanolamine pH 8.5. The surface was equilibrated with running buffer composed of 20 mM HEPES 150 mM NaCl, pH 7.2, 3 mM EDTA and 0.005% Tween 20. The CD80-Ig and CD80-Ig α tp (20 ug/ml) diluted in running buffer were injected over the surface (60 ul)

25 with a flowrate of 10 ul/min. The results show that CD80-Ig dissociates very rapidly from the CD28-Ig coated surface, whereas the rate of dissociation for CD80-Ig α tp is about three orders of magnitude slower (Figure 14).

For comparison of CD80-Ig α tp and CD86-Ig α tp binding with immobilized CD28-Ig, solutions of CD80-Ig α tp and CD86-Ig α tp (5 ug/ml) were prepared in running buffer

30 described above. Sample solutions were injected (60 ul) at 10 ul/min. Between samples, the surface was regenerated with a 30 ul injection of Gentle elution buffer (Pierce Chemicals, Rockville, ILL). The results show that like CD80-Ig α tp above, CD86-Ig α tp dissociates slowly from the CD28-Ig surface (Figure 15). The off-rate for CD86-Ig α tp

appears higher than that for the CD80 construct, consistent with the weaker binding of this protein in the ELISA (part 1, above) and cell binding (part 3, below) formats and the lower intrinsic affinity measured by calorimetry (part 4, below).

3) Binding to cells expressing cell-surface CD28

- 5 By flow cytometry, both CD80- and CD86-Ig α tp show specific binding to CD28 positive cells (Figs. 16A and 16B), whereas no binding is observed with CD80- or CD86-Ig themselves (not shown). Non-adherent CD28 expressing cells (PCD28.1.s2.1) were used for this assay format. PCD28.1.S2.1 cells were created by transfection of PE30.2 cells [D. Emilie *et al.*, Eur. J. Immunol., **19**:1619-1624 (1989)] with a vector for expression of human
- 10 CD28 [D. Couez *et al.*, Molecular Immunology, **31**:47-57 (1994)]. All incubations were carried out on ice. Unlabelled CD80- and CD86-Ig α tp or CD80- and CD86-Ig were incubated with the cells in binding buffer consisting of PBS w/o Ca+2/Mg+2, 0.2% bovine serum albumin and 0.1% sodium azide. After washing twice in binding buffer, bound fusion proteins were detected with a 1:2000 dilution of a goat anti-human polyclonal
- 15 antibody labeled with FITC (Flourescein isothiocyanate, Southern Biotech). Following two additional washes in binding buffer, the cells were resuspended in binding buffer and analyzed on a FACSsort analyzer (Becton-Dickinson) using a 488nm laser. Low non-specific binding was shown by incubating the cells with a 200 fold excess of a CD28 monoclonal antibody 28.1 prior to ("block" curves, Figure C), or at the same time as
- 20 ("competition" curves, Figs. 16A and 16B), or by addition of CTLA4-Ig with the CD80 and CD86 fusion proteins (not shown).

4) Binding to CD28-Ig and CTLA4-Ig in solution

- The solution binding properties of the hexameric and standard Ig fusion proteins of CD80 and CD86 were compared using isothermal titration calorimetry, essentially as described
- 25 previously for other proteins [K. Johanson *et. al.*, J. Biol. Chem., **270**: 9459-9471 (1995), and references therein]. The binding to CTLA4-Ig is summarized in the following table:

Table I

Direct comparison of CTLA4-Ig binding properties of CD80 and CD86 Ig fusion proteins
as measured by titration calorimetry at 2 temperatures

5

Construct	Experimental Temperature	K_d , nM at Temperature	Molar Ratio (CTLA4-Ig per construct)	ΔH , kcal/mol	K_d nM at 37°C
CD80-Ig	37°C	5.4	0.72	-32 ± 2	5
CD80-Ig α tp	37°C	5.9	4.02	-36 ± 2	6
CD86-Ig	37°C	38	0.80	-33 ± 3	40
CD86-Ig α tp	37°C	20	2.72	-37 ± 4	20
CD80-Ig	44°C	4.9	0.60	-36 ± 4	2
CD80-Ig α tp	44°C	8.1	3.79	-38 ± 3	2
CD86-Ig	44°C	71	0.79	-34 ± 4	22
CD86-Ig α tp	44°C	38	2.83	-41 ± 4	9

- The error in K_d 's is about a factor of 2 and the error in molar binding ratio's is 10-20%. K_d values at 37°C were either measured directly at 37°C or were corrected for temperature differences using the van't Hoff equation, as described in M. L. Doyle et. al., J. Mol. Recognition, 2: 65-74 (1996). Concentrations were defined by absorbance at 280 nm using the following: A) molecular masses of 90,059 (CTLA4-Ig), 127,000 (CD80-Ig), 810,000 (CD80-Ig α tp), 140,000 (CD86-Ig), and 890,000 (CD86-Ig α tp) and B) calculated extinction coefficients of 1.22 (CTLA4-Ig), 1.10 (CD80-Ig and CD80-Ig α tp), and 1.03 (CD86-Ig and CD86-Ig α tp). The molecular weights for CTLA4-Ig, CD80-Ig, and CD86-Ig were determined by mass spectral analysis. The molecular masses of CD80-Ig α tp and CD86-Ig α tp were estimated as 6x the mass of the respective Ig proteins plus 40,000 Da. contributed by the twelve tailpiece segments.

Direct comparison of the CTLA4-Ig binding to CD80-Ig, CD80-Ig α tp, CD86-Ig, and CD86-Ig α tp constructs in solution phase by isothermal titration calorimetry

d monstrates several features. First, the affinities of the Ig versus Ig- α tp constructs are equivalent in solution. This suggests that, as expected, solution binding affinities of the α tp constructs do not benefit from avidity effects like they do in ELISA and cell binding assays. Second, the enthalpy changes which accompany the molecular interactions of the Ig and Ig α tp constructs are also the same and support the view that the molecular details of the interactions are the same. Third, the titration equivalence points for CTLA4-Ig binding to the CD80 and CD86 Ig versus Ig α tp constructs indicate that all these reagents were $\geq 50\%$ active during the calorimetry assay. With regard to this latter point, comparison of the Ig α tp and Ig constructs shows a ratio of about 6 for CD80, indicating about equivalent binding activity for the CD80 domains in both constructs. The lower ratio for the corresponding CD80-Ig α tp protein indicates some loss of activity in this preparation.

Interactions of CD28-Ig with either CD80- or CD86-Ig were not detected in solution by calorimetry, suggesting an affinity of interaction weaker than 1 μ M. This lower affinity for CD28 than for CTLA4 is in agreement with other reports [P. S. Linsley et. al., *Immunity* 1: 793-801 (1994)]. CD28-Ig also did not show detectable binding to CD80- or CD86-Ig α tp, which is consistent with the solution affinities of the α tp constructs not benefiting from avidity effects.

Example 2 - Demonstration of agonist activity for the CD80- and CD86 Ig α tp protein

20 A. CTLL-2 bioassay for detection of IL-2 levels

The CD80- and CD86-Ig α tp proteins were compared to the corresponding CD80 and CD86-Ig proteins to determine their ability to stimulate cells expressing human CD28 using two murine T-cell hybridoma cell lines expressing human CD28, PCD28.1.s2.1 and DCL27CD28wt.s2. The PCD28.1.s2.1 cell line was described in Example 1, part 3. The DCL27CD28wt.s2 cell line was created by transfection of the DC27 cell line [F. Pages et al., *Nature*, 369:327-329 (1994); F. Pages et al., *J. Biol. Chem.*, 271(16):9403-9409 (1996)] with the same CD28 expression vector used for the PCD28.1.s2.1 cells [D. Couez et al., *Molecular Immunology*, 31:47-57 (1994)]. These cell lines were examined for their ability to produce IL-2 in response to activation with CD80- and CD86-Ig in comparison with the corresponding Ig α tp fusion proteins. 96-well plates were coated with or without a CD3 antibody together with the CD80 and CD86 fusion proteins. This was accomplished by first incubating the plates with a previously determined suboptimal concentration of hamster anti-human CD3 antibody (MAb 145-2C11, Boehringer-Mannheim Biochemicals) for two

hours at room temperature (RT) or with just buffer alone, washing the plates with PBS, adding goat anti-human Ig heavy chain (GAH-IgHc, Sigma Chemical Co.) for an additional two hours at RT, washing again and coating with different concentrations of the CD80 or CD86 fusion protein for 16-18 hours at 4°C, washing again, and finally blocking for 30 min. with 0.2% BSA-PBS. T cells (1×10^5 /well) were added in 150 μ l medium into duplicate wells. For comparison, the soluble fusion proteins and the 248.23.2 CD28 MAb (IgM) [A. Morretta, University of Genova, Italy] were added to non-coated wells. T cells were incubated in the wells for 24 hours at 37°C, and supernatants were collected and evaluated for IL-2 levels in a standard CTLL-2 bioassay [S.M. Gillis *et al.*, *J. Immunol.*, **120**:2027 (1978)]. Briefly, 1×10^4 IL-2 dependent CTLL-2 cells (ATCC)/well in 75 μ l medium were added to an equal volume of test supernatant and incubated for 24 hours at 37°C. The cells were pulsed with 10 μ l of 5 mg/ml MTT (Sigma Chemical Co.) for 4 hours, and lysed with 100 μ l 10% SDS/0.01N HCl solution for 14-16 hours. OD₅₇₀ readings were converted into ng/ml of IL-2 based on a standard curve generated by treating cells with known concentrations of IL-2.

In all assays, the CD80- and CD86-Ig α p proteins were more efficient stimulators of the CD28 T-cells than the corresponding monomeric Ig constructs (Figs. 17 and 18). The soluble hexameric proteins induced IL-2 production in the absence of CD3 crosslinking (GAH), whereas under the same conditions, no activity was observed with CD80- or CD86-Ig themselves. A similar level of IL-2 induction was observed with the oligomeric CD28 IgM antibody 248.23.2. Cross-linking of the hexameric CD80 and CD86 proteins with GAH antibody increased the IL-2 response relative to the absence of cross-linker, but still did not give a response for the monomeric Ig constructs. In the presence of CD3 antibody, the differences between the hexameric and monomeric Ig fusion proteins were minimal, being about 2-fold or less.

B. Fluorescein di-b-D-galactosidase (FDG) assay for detection of IL-2 promoter activity

A second assay for agonist activity measured induction of IL-2 promoter activity, rather than production of IL-2 protein. The PCD28.1.S2.1 cells described above also contain *lacZ* fused to the IL-2 promoter. Thus, the PCD28.1.S2.1 cell line provides a convenient system for measuring IL-2 promoter activity upon CD28-mediated T cell simulation. T cells were activated as described above for the CTLL-2 assay, spun down, resuspended in 50 μ l of media + 50 μ l of PBS, lysed with 10 μ l of 20% Triton X-100, and

supplemented with 25 μ l of 10 mM FDG (Molecular Probes), a fluorogenic substrate for b-galactosidase. Hydrolysis of FDG first yields fluorescein monogalactoside (FMG) and then the highly fluorescent product fluorescein. Cell lysates were incubated with FDG for 60 min., and the levels of fluorescence were measured by Fluoroscan (MTX Lab Systems, Inc).

5 The results of these assays (Fig. 19) were similar to those described above for IL-2 production. The primary difference was that low levels of IL-2 promoter activity were observed for the monomeric Ig proteins.

C. Stimulation of CD28 cells by CD80- and CD86-Ig α p proteins in solution

In the above examples (parts A and B), the Ig and Ig α p proteins showed the
10 greatest activity when captured on the surface of the microtiter well.. However, the CD80- and CD86-Ig α p proteins were also able to stimulate CD28 cells when added directly to the cells in solution, whereas no response was observed with the corresponding standard Ig fusion proteins. CD80- and CD86-Ig α p showed a dose dependent stimulation of IL-2 promoter activity (Fig 20A) and IL-2 production (Fig 20B) when added to PC28.1.s2.1
15 cells. In contrast, no stimulation was observed with CD80-Ig (Figs. 20A and 20B) or CD86-Ig (not shown). IL-2 promoter activity and IL-2 levels were measured similarly to that described in parts A and B above, except that proliferation of the reader CTLL-2 cells was measured by ³H-thymidine incorporation. The level of response at near saturation levels of CD80- and CD86-Ig α p proteins (1 μ g/ml) was comparable to that observed for
20 stimulation through cross-linking of CD3 with immobilized CD3 antibody (Fig 20C). The specificity of the response to CD80- and CD86-Ig α p was confirmed by complete blockade with the addition of CTLA4-Ig (not shown).

In summary, the results from these assays show that the CD80- and CD86- Ig α p proteins have agonist activity under conditions where little or no activity was observed for the
25 corresponding monomeric Ig proteins.

Example 3 - Compound screen assay for identifying small molecule antagonists of the interaction between CD28 and CD80

An ELISA format was used to identify small molecule antagonists of CD80 and
30 CD86 binding to CD28 by screening a large bank of chemical compounds and natural products. The assay was carried out as in the format described in Example 1, part E.1, except that immediately following addition of the biotinylated CD80-Ig α p (222 ng/ml in a volume of 90 μ l), dilutions of test compound were added (10 μ l). The compounds were

dissolved at 100x assay concentration in dimethyl sulfoxide (DMSO) and subsequently diluted in 50%DMSO/50% H₂O to a 10X working stock. The assay was not sensitive (<10% alteration of signal) to DMSO at concentrations of 5% or less.

Results from one test assay are summarized in Fig. 21 and Table II. The BM-34 test set consists of 968 compounds in two formats - as individual compounds and as 88 multimixes with 11 individual compounds in each multimix sample. Both BM-34 formats were assayed (at a concentration of 200 µg/ml for each multimix sample and 20 µg/ml for individual compounds) for inhibition of biotinylated CD80-Igαtp binding to immobilized CD28-Ig in 96 well plates. Results for setting a 70% or 85% cutoff for inhibition are shown in Table II. In Fig. 21, the percent inhibition range is plotted against the number of compounds showing the indicated range of inhibition. The low percentage of compounds showing activity in the 80-90 % range of inhibition makes this a suitable threshold for rapid screening.

Table II
CD80-Igαtp Screen Assay Results
BM-34 Test Compound Set

<u>Result</u>	<u>70% Cutoff</u>	<u>85% Cutoff</u>
Hits on Multimix plates	7	1
+ Multimix Samples with + Compound	5/7	1/1
- Multimix Samples with + Compound	8	0

As illustrated in this table, eight of the mixes gave 70% or greater inhibition. Deconvolution by assay of the individual compounds from these 8 mixes at 20 µg/ml confirmed that there was a compound with comparable activity. The two other mixes that failed to confirm had one or more compounds with activity very close to the 70% inhibition

observed in the original multimix assay. Selecting a higher cutoff of 85% gave only one multimix hit and that was confirmed in the assay of individual compounds. Further evidence of reproducibility and selectivity was that only eight compounds from mixes below the 70% cutoff showed > 70% inhibition when assayed individually. Selectivity was further increased by reducing the concentration of the multimix samples to 100 µg/ml and the individual compounds to 10 µg/ml. This corresponds to about a 30 µM concentration for the compounds since their average MW is 300-400 daltons. At 100 µg/ml, multimix samples showed a desired shift to lower average inhibition (90% of the mixtures gave 60% or less inhibition) while retaining an acceptable hit rate at a high level of inhibition (4% of the mixtures giving 70% or greater inhibition). Through the use of this assay, small molecule inhibitors of the interaction of CD80 with CD28 can be identified.

Example 4 - α p-mediated oligomerization of a mouse/human IgG1 chimeric antibody.

To examine the generalization of α p-mediated hexamer formation of the Fc region of human IgG, the α p segment was introduced into a chimeric antibody containing heavy and light chain variable regions from the mouse monoclonal antibody 1C8 and the human kappa and IgG1 constant regions. 1C8 is directed against the human EPO (erythropoietin) receptor. The α p sequence was introduced onto the heavy chain of the antibody by replacing the Eco RI/Sac II fragment of CD80Fc α tplink with the Eco RI/Sac II fragment of EpoR(CH)IgG1-PCN, a vector containing the heavy chain of the chimeric 1C8 antibody, to give the vector EpoR(CH)Fc α tplink. In both vectors, Eco RI cleaves between the CMV promoter and the start of the N-terminal signal sequences and Sac II cleaves at a conserved site in constant region 2 of the human heavy chain.

Test samples of the hexameric mAb were produced in COS-7 cells upon co-transfection of EpoR(CH)Fc α tplink and a vector for expression of the light chimeric light chain, following procedures described above in Example 1, part C. Initially, 5 T150 flasks were co-transfected with the two vectors and 300 ml of conditioned media were collected. The hexameric antibody was purified by affinity chromatography on Protein A. Purity was about 90% as determined by coomassie staining of the sample as analyzed by reducing SDS/PAGE. Under nonreducing conditions on SDS/PAGE, the antibody migrated in the size range of IgM (not shown).

The sample was further characterized by analytical size exclusion chromatography on a 3.2 X 30 mm Superose 6 column run on a Smart System HPLC (Pharmacia Biotech,

Piscataway NJ). The major peak (Fig. 22) corresponds to binding activity, as monitored in an ELISA using a recombinant human EPO receptor Ig fusion protein (EPOr-Ig), and eluted at a size consistent with hexamer formation (anti-EPOr-IgG₁ α p). The parental chimeric antibody (anti-EPOr-IgG₁) elutes substantially later from the column and is represented in the figure by the dashed lines.

These results indicate that addition of the α p segment to the human IgG1 constant region leads to formation of hexameric antibody.

Numerous modifications and variations of the present invention are included in the above-identified specification and are expected to be obvious to one of skill in the art. Such modifications and alterations to the compositions and processes of the present invention are believed to be encompassed in the scope of the claims appended hereto.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: Sweet, Raymond W.
Truneh, Alemseged
Chaikin, Margery A.
Lyn, Sally D.P.

(ii) TITLE OF INVENTION: Hexameric Fusion Proteins and Uses
Therefor

(iii) NUMBER OF SEQUENCES: 27

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: SmithKline Beecham Corporation -
Corporate Patents
(B) STREET: 709 Swedeland Road
(C) CITY: King of Prussia
(D) STATE: Pennsylvania
(E) COUNTRY: USA
(F) ZIP: 19406-2799

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: US
(B) FILING DATE: 13-June-1997
(C) CLASSIFICATION:

PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: US
(B) FILING DATE: 14-June-1996
(C) CLASSIFICATION: 60/019,934

PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: US
(B) FILING DATE: 19-FEB-1997
(C) CLASSIFICATION: 60/043,948

PRIOR APPLICATION DATA:

- (A) APPLICATION NUMBER: US
- (B) FILING DATE: 21-FEB-1997
- (C) CLASSIFICATION: 60/038,915

(viii) ATTORNEY/AGENT INFORMATION:

- (A) NAME: Gimmi, Edward R.
- (B) REGISTRATION NUMBER: 38,891
- (C) REFERENCE/DOCKET NUMBER: P50496

(ix) TELECOMMUNICATION INFORMATION:

- (A) TELEPHONE: 610-270-5015
- (B) TELEFAX: 610-270-5090

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 7167 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

GACGTCGACG GATCGGGAGA TCGGGGATCG ATCCGTCGAC GTACGACTAG TTATTAATAG	60
TAATCAATTA CGGGGTCATT AGTTCATAGC CCATATATGG AGTTCGCGCT TACATAACTT	120
ACGGTAAATG GCCCGCCTGG CTGACCGCCC AACGACCCCC GCCCATGAC GTCAATAATG	180
ACGTATGTTT CCATAGTAAC GCCAATAGGG ACTTTCCATT GACGTCAATG GGTGGACTAT	240
TTACGGTAAA CTGCCCCTT GGCAGTACAT CAAGTGTATC ATATGCCAAG TACGCCCCCT	300
ATTGACGTCA ATGACGGTAA ATGGCCCGCC TGGCATTATG CCCAGTACAT GACCTTATGG	360
GACTTTCCTA CTTGGCAGTA CATCTACGTA TTAGTCATCG CTATTACCAT GGTGATGCGG	420
TTTTGGCAGT ACATCAATGG GCGTGGATAG CGGTTTGA CTACGGGGATT TCCAAGTCTC	480
CACCCCATG ACGTCAATGG GAGTTTGT TTGGCACCAA ATCAACGGGA CTTTCCAAAA	540

TGTCGTAACA ACTCCGCCCC ATTGACGCAA ATGGGCGGTA GGCCTGTACG GTGGGAGGTC	600
TATATAAGCA GAGCTGGGTA CGTGAACCGT CAGATCGCCT GGAGACGCCA TCGAATTCGG	660
TTACCTGCAG ATATCAAGCT AATTCGGTAC CGGGCCCCCC TCGAGCCTGA AGCCATGGGC	720
CACACACGGA GGCAGGGAAC ATCACCATCC AAGTGTCCAT ACCTCAATTT CTTTCAGCTC	780
TTGGTGCTGG CTGGTCTTTC TCACTTCTGT TCAGGTGTTA TCCACGTGAC CAAGGAAGTG	840
AAAGAAGTGG CAACGCTGTC CTGTGGTCAC AATGTTTCTG TTGAAGAGCT GGCACAAACT	900
CGCATCTACT GGCAAAGGA GAAGAAAATG GTGCTGACTA TGATGTCTGG GGACATGAAT	960
ATATGGCCCG AGTACAAGAA CCGGACCATC TTTGATATCA CTAATAACCT CTCCATTGTG	1020
ATCCTGGCTC TGGGCCCATC TGACGAGGGC ACATACGAGT GTGTTGTTCT GAAGTATGAA	1080
AAAGACGCTT TCAAGCGGGA ACACCTGGCT GAAGTGACGT TATCAGTCAA AGCTGACTTC	1140
CCTACACCTA GTATATCTGA CTTTGAAATF CCAACTTCTA ATATTAGAAG GATAATTTGC	1200
TCAACCTCTG GAGGTTTTC AGAGCCTCAC CTCTCCTGGT TGGAAAATGG AGAAGAATTA	1260
AATGCCATCA ACACAACAGT TTCCCAAGAT CCTGAACTG AGCTCTATGC TGTTAGCAGC	1320
AAACTGGATT TCAATATGAC AACCAACCAC AGCTTCATGT GTCTCATCAA GTATGGACAT	1380
TTAAGAGTGA ATCAGACCTT CAACTGGAAT ACAACCAAGC AAGAGCATTT TCCTGATCAG	1440
GAGCCCAAAT CGGCCGACAA AACTCACACA TGCCCAACCGT GCCCAGCACC TGAACCTCTG	1500
GGGGGACCGT CAGTCTTCCT CTCCCCCA AAACCAAGG ACACCCTCAT GATCTCCCGG	1560
ACCCCTGAGG TCACATGCGT GGTGGTGGAC GTGAGCCACG AAGACCCTGA GGTCAAGTTC	1620
AACTGGTACG TGGACGGCGT GGAGGTGCAT AATGCCAAGA CAAAGCCGCG GGAGGAGCAG	1680
TACAACAGCA CGTACCGGGT GGTGAGCGTC CTCACCGTCC TGCACCAGGA CTGGCTGAAT	1740
GGCAAGGAGT ACAAGTGCAA GGTCTCCAAC AAAGCCCTCC CAGCCCCCAT CGAGAAAACC	1800
ATCTCCAAAG CCAAAGGGCA GCCCCGAGAA CCACAGGTGT ACACCCTGCC CCCATCCCGG	1860

GATGAGCTGA CCAAGAACCA GGTCAGCCTG ACCTGCCTGG TCAAAGGCTT CTATCCCAGC	1920
GACATCGCCG TGGAGTGGGA GAGCAATGGG CAGCCGGAGA ACAACTACAA GACCACGCCT	1980
CCCGTGCTGG ACTCCGACGG CTCCTTCTTC CTCTACAGCA AGCTCACCGT GGACAAGAGC	2040
AGGTGGCAGC AGGGGAACGT CTTCTCATGC TCCGTGATGC ATGAGGCTCT GCACAACCAC	2100
TACACGCAGA AGAGCCTAAG CTTGTCTGCG GGTAACCCA CCCATGTCAA TGTGTCTGTT	2160
GTCATGGCGG AGGTGGACGG CACCTGCTAC TGATAGTCTA GAGCTCGCTG ATCAGCCTCG	2220
ACTGTGCCTT CTAGTTGCCA GCCATCTGTT GTTTGCCCCCT CCCCCGTGCC TTCCTTGACC	2280
CTGGAAGGTG CCACTCCCAC TGTCTTTCC TAATAAAATG AGGAAATTGC ATCGCATTGT	2340
CTGAGTAGCT GTCATTCTAT TCTGGGGGGT GGGGTGGGGC AGGACAGCAA GGGGGAGGAT	2400
TGGGAAGACA ATAGCAGGCA TGCTGGGGAT GCGGTGGGCT CTATGGAACC AGCTGGGGCT	2460
CGAGGGGGGA TCTCCCGATC CCCAGCTTTG CTTCTCAATT TCTTATTTGC ATAATGAGAA	2520
AAAAAGGAAA ATTAATTTTA ACACCAATTC AGTAGTTGAT TGAGCAAATG CGTTGCCAAA	2580
AAGGATGCTT TAGAGACAGT GTTCTCTGCA CAGATAAGGA CAAACATTAT TCAGAGGGAG	2640
TACCCAGAGC TGAGACTCCT AAGCCAGTGA GTGGCACAGC ATTCTAGGGA GAAATATGCT	2700
TGTCATCACC GAAGCCTGAT TCCGTAGAGC CACACCTTGG TAAGGGCCAA TCTGCTCACA	2760
CAGGATAGAG AGGGCAGGAG CCAGGGCAGA GCATATAAGG TGAGGTAGGA TCAGTTGCTC	2820
CTCACATTTC CTTCTGACAT AGTTGTGTTG GGAGCTTGGG TAGCTTGGAC AGCTCAGGGC	2880
TGCGATTTTC CGCCAACTT GACGGCAATC CTAGCGTGAA GGCTGGTAGG ATTTTATCCC	2940
CGCTGCCATC ATGGTTCGAC CATTGAACTG CATCGTCGCC GTGTCCCAA ATATGGGGAT	3000
TGGCAAGAAC GGAGACCTAC CCTGGCCTCC GCTCAGGAAC GAGTTCAAGT ACTTCCAAAG	3060
AATGACCACA ACCTCTTCAG TGGAAGGTAA ACAGAATCTG GTGATTATGG GTAGGAAAAC	3120
CTGGTTCTCC ATTCCTGAGA AGAATCGACC TTAAAGGAC AGAATTAATA TAGTTCTCAG	3180

TAGAGAACTC AAAGAACCAC CACGAGGAGC TCATTTTCTT GCCAAAAGTT TGGATGATGC	3240
CTTAAGACTT ATTGAACAAC CGGAATTGGC AAGTAAAGTA GACATGGTTT GGATAGTCGG	3300
AGGCAGTTCT GTTTACCAGG AAGCCATGAA TCAACCAGGC CACCTTAGAC TCTTTGTGAC	3360
AAGGATCATG CAGGAATTG AAAGTGACAC GTTTTCCCA GAAATTGATT TGGGGAAATA	3420
TAAACTTCTC CCAGAATACC CAGGCGTCCT CTCTGAGGTC CAGGAGGAAA AAGGCATCAA	3480
GTATAAGTTT GAAGTCTACG AGAAGAAAGA CTAACAGGAA GATGCTTTCA AGTTCTCTGC	3540
TCCCCTCCTA AAGCTATGCA TTTTATAAG ACCATGCTAG CTTGAAGTTG TTTATTGCAG	3600
CTTATAATGG TTACAAATAA AGCAATAGCA TCACAAATTT CACAAATAAA GCATTTTTTT	3660
CACTGCATTC TAGTTGTGGT TTGTCCAAAC TCATCAATGT ATCTTATCAT GTCTGGATCA	3720
ACGATAGCTT ATCTGTGGGC GATGCCAAGC ACCTGGATGC TGTGGTTTC CTGCTACTGA	3780
TTTAGAAGCC ATTTGCCCCC TGAGTGGGGC TTGGGAGCAC TAACTTTCTC TTTCAAAGGA	3840
AGCAATGCAG AAAGAAAAGC ATACAAAGTA TAAGCTGCCA TGTAATAATG GAAGAAGATA	3900
AGGTTGTATG AATTAGATTT ACATACTTCT GAATTGAAAC TAAACACCTT TAAATTCTTA	3960
AATATATAAC ACATTTTATA TGAAAGTATT TTACATAAGT AACTCAGATA CATAGAAAAC	4020
AAAGCTAATG ATAGGTGTCC CTAAAAGTTC ATTTATTAAT TCTACAAATG ATGAGCTGGC	4080
CATCAAAATT CCAGCTCAAT TCTTCAACGA ATTAGAAAGA GCAATCTGCA AACTCATCTG	4140
GAATAACAAA AAACCTAGGA TAGCAAAAAC TCTTCTCAAG GATAAAAGAA CCTCTGGTGG	4200
AATCACCATG CCTGACCTAA AGCTGTACTA CAGAGCAATT GTGATAAAAA CTGCATGGTA	4260
CTGATATAGA AACGGACAAG TAGACCAATG GAATAGAACC CACACACCTA TGGTCACTTG	4320
ATCTTCAACA AGAGAGCTAA AACCATCCAC TGGAAAAAG ACAGCATTTT CAACAAATGG	4380
TGCTGGCACA ACTGGTGGTT ATCATGGAGA AGAATGTGAA TTGATCCATT CCAATCTCCT	4440
TGTACTAAGG TCAAATCTAA GTGGATCAAG GAACTCCACA TAAAACCAGA GAACTGAAA	4500

CTTATAGAGG AGAAAGTGGG GAAAAGCCTC GAAGATATGG GCACAGGGGA AAAATTCCTG	4560
AATAGAACAG CAATGGCTTG TGCTGTAAGA TCGAGAATTG ACAAATGGGA CCTCATGAAA	4620
CTCCAAAGCT ATCGGATCAA TTCCTCCAAA AAAGCCTCCT CACTACTTCT GGAATAGCTC	4680
AGAGGCCGAG GCGGCCTCGG CCTCTGCATA AATAAAAAAA ATTAGTCAGC CATGCATGGG	4740
GCGGAGAATG GCGGGAATG GCGGAGTTA GGGGCGGGAT GGGCGGAGTT AGGGGCGGGA	4800
CTATGGTTGC TGAATAATTG AGATGCATGC TTTGCATACT TCTGCCTGCT GGGGAGCCTG	4860
GGGACTTTCC ACACCTGGTT GCTGACTAAT TGAGATGCAT GCTTTGCATA CTTCTGCCTG	4920
CTGGGGAGCC TGGGGACTTT CCACACCCTA ACTGACACAC ATTCCACAGA ATTAATTCCC	4980
GATCCCGTCG ACCTCGAGAG CTTGGCGTAA TCATGGTCAT AGCTGTTTCC TGTGTGAAAT	5040
TGTTATCCGC TCACAATTCC ACACAACATA CGAGCCGGAA GCATAAAGTG TAAAGCCTGG	5100
GGTGCCTAAT GAGTGAGCTA ACTCACATTA ATTGCGTTGC GCTCACTGCC CGCTTTCCAG	5160
TCGGGAAACC TGTCGTGCCA GCTGCATTAA TGAATCGGCC AACGCGCGGG GAGAGGCGGT	5220
TTGCGTATTG GCGCTCTTC CGCTTCCTCG CTCACTGACT CGCTGCGCTC GGTGCTTCGG	5280
CTGCGGCGAG CGGTATCAGC TCACTCAAAG GCGTAATAC GGTATCCAC AGAATCAGGG	5340
GATAACGCAG GAAAGAACAT GTGAGCAAAA GGCCAGCAAA AGGCCAGGAA CCGTAAAAAG	5400
GCCGCGTTGC TGGCGTTTTT CCATAGGCTC CGCCCCCTG ACGAGCATCA CAAAAATCGA	5460
CGCTCAAGTC AGAGGTGGCG AAACCCGACA GGAATAAAA GATACCAGGC GTTTCCCCCT	5520
GGAAGCTCCC TCGTGCGCTC TCCTGTTCCG ACCCTGCCGC TTACCGGATA CCTGTCCGCC	5580
TTTCTCCCTT CGGGAAGCGT GCGCTTTCT CAATGCTCAC GCTGTAGGTA TCTCAGTTCTG	5640
GTGTAGGTCG TTCGCTCCAA GCTGGGCTGT GTGCACGAAC CCCCCGTTCA GCCCGACCGC	5700
TGCGCCTTAT CCGGTAACATA TCGTCTTGAG TCCAACCCGG TAAGACACGA CTTATCGCCA	5760
CTGGCAGCAG CCACTGGTAA CAGGATTAGC AGAGCGAGGT ATGTAGGCGG TGCTACAGAG	5820

TTCTTGAAGT GGTGGCCTAA CTACGGCTAC ACTAGAAGGA CAGTATTTGG TATCTGCGCT	5880
CTGCTGAAGC CAGTTACCTT CGGAAAAAGA GTTGGTAGCT CTTGATCCGG CAAACAAACC	5940
ACCGCTGGTA GCGGTGGTTT TTTTGTTTGC AAGCAGCAGA TTACGCGCAG AAAAAAAGGA	6000
TCTCAAGAAG ATCCTTTGAT CTTTCTACG GGGTCTGACG CTCAGTGGAA CGAAAACTCA	6060
CGTTAAGGGA TTTTGGTCAT GAGATTATCA AAAAGGATCT TCACCTAGAT CCTTTTAAAT	6120
TAAAAATGAA GTTTTAAATC AATCTAAAGT ATATATGAGT AAACCTGGTC TGACAGTTAC	6180
CAATGCTTAA TCAGTGAGGC ACCTATCTCA GCGATCTGTC TATTTGTTTC ATCCATAGTT	6240
GCCTGACTCC CCGTCGTGTA GATAACTACG ATACGGGAGG GCTTACCATC TGGCCCCAGT	6300
GCTGCAATGA TACCGCGAGA CCCACGCTCA CCGGCTCCAG ATTTATCAGC AATAAACCAG	6360
CCAGCCGGAA GGGCCGAGCG CAGAAGTGGT CCTGCAACTT TATCCGCCTC CATCCAGTCT	6420
ATTAATTGTT GCCGGGAAGC TAGAGTAAGT AGTTCGCCAG TTAATAGTTT GCGCAACGTT	6480
GTTGCCATTG CTACAGGCAT CGTGGTGTC CGCTCGTCGT TTGGTATGGC TTCATTACAGC	6540
TCCGGTCCC AACGATCAAG GCGAGTTACA TGATCCCCCA TGTGTGCAA AAAAGCGGTT	6600
AGCTCCTTCG GTCCTCCGAT CGTTGTCAGA AGTAAGTTGG CCGCAGTGTT ATCACTCATG	6660
GTTATGGCAG CACTGCATAA TTCTCTTACT GTCATGCCAT CCGTAAGATG CTTTTCTGTG	6720
ACTGGTGAGT ACTCAACCAA GTCATTCTGA GAATAGTGTA TCGGCGACC GAGTTGCTCT	6780
TGCCCCGGGT CAATACGGGA TAATACCGCG CCACATAGCA GAACTTTAAA AGTGCTCATC	6840
ATTGGAACAC GTTCTTCGGG GCGAAAACTC TCAAGGATCT TACCGCTGTT GAGATCCAGT	6900
TCGATGTAAC CCACTCGTGC ACCCAACTGA TCTTCAGCAT CTTTACTTT CACCAGCGTT	6960
TCTGGGTGAG CAAAAACAGG AAGGCAAAAT GCCGCAAAA AGGGAATAAG GGCGACACGG	7020
AAATGTTGAA TACTCATACT CTCCTTTTT CAATATTATT GAAGCATTTA TCAGGGTTAT	7080
TGTCTCATGA GCGGATACAT ATTTGAATGT ATTTAGAAAA ATAAACAAAT AGGGGTCCG	7140

CGCACATTTC CCCGAAAAGT GCCACCT

7167

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1560 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 63..1538

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

GAATTCGGTT ACCTGCAGAT ATCAAGCTAA TTCGGTACCG GGCCCCCCTC GAGCCTGAAG	60
CC ATG GGC CAC ACA CGG AGG CAG GGA ACA TCA CCA TCC AAG TGT CCA	107
Met Gly His Thr Arg Arg Gln Gly Thr Ser Pro Ser Lys Cys Pro	
1 5 10 15	
TAC CTC AAT TTC TTT CAG CTC TTG GTG CTG GCT GGT CTT TCT CAC TTC	155
Tyr Leu Asn Phe Phe Gln Leu Leu Val Leu Ala Gly Leu Ser His Phe	
20 25 30	
TGT TCA GGT GTT ATC CAC GTG ACC AAG GAA GTG AAA GAA GTG GCA ACG	203
Cys Ser Gly Val Ile His Val Thr Lys Glu Val Lys Glu Val Ala Thr	
35 40 45	
CTG TCC TCT GGT CAC AAT GTT TCT GTT GAA GAG CTG GCA CAA ACT CGC	251
Leu Ser Cys Gly His Asn Val Ser Val Glu Glu Leu Ala Gln Thr Arg	
50 55 60	
ATC TAC TGG CAA AAG GAG AAG AAA ATG GTG CTG ACT ATG ATG TCT GGG	299
Ile Tyr Trp Gln Lys Glu Lys Lys Met Val Leu Thr Met Met Ser Gly	
65 70 75	

PCT/US97/12599

44

TGC CCA GCA CCT GAA CTC CTG GGG GGA CCG TCA GTC TTC CTC TTC CCC	875
Cys Pro Ala Pro Glu Leu Leu Gly Gly Pro Ser Val Phe Leu Phe Pro	
260 265 270	
CCA AAA CCC AAG GAC ACC CTC ATG ATC TCC CGG ACC CCT GAG GTC ACA	923
Pro Lys Pro Lys Asp Thr Leu Met Ile Ser Arg Thr Pro Glu Val Thr	
275 280 285	
TGC GTG GTG GTG GAC GTG AGC CAC GAA GAC CCT GAG GTC AAG TTC AAC	971
Cys Val Val Val Asp Val Ser His Glu Asp Pro Glu Val Lys Phe Asn	
290 295 300	
TGG TAC GTG GAC GGC GTG GAG GTG CAT AAT GCC AAG ACA AAG CCG CGG	1019
Trp Tyr Val Asp Gly Val Glu Val His Asn Ala Lys Thr Lys Pro Arg	
305 310 315	
GAG GAG CAG TAC AAC AGC ACG TAC CGG GTG GTC AGC GTC CTC ACC GTC	1067
Glu Glu Gln Tyr Asn Ser Thr Tyr Arg Val Val Ser Val Leu Thr Val	
320 325 330 335	
CTG CAC CAG GAC TGG CTG AAT GGC AAG GAG TAC AAG TGC AAG GTC TCC	1115
Leu His Gln Asp Trp Leu Asn Gly Lys Glu Tyr Lys Cys Lys Val Ser	
340 345 350	
AAC AAA GCC CTC CCA GCC CCC ATC GAG AAA ACC ATC TCC AAA GCC AAA	1163
Asn Lys Ala Leu Pro Ala Pro Ile Glu Lys Thr Ile Ser Lys Ala Lys	
355 360 365	
GGG CAG CCC CGA GAA CCA CAG GTG TAC ACC CTG CCC CCA TCC CGG GAT	1211
Gly Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser Arg Asp	
370 375 380	
GAG CTG ACC AAG AAC CAG GTC AGC CTG ACC TGC CTG GTC AAA GGC TTC	1259
Glu Leu Thr Lys Asn Gln Val Ser Leu Thr Cys Leu Val Lys Gly Phe	
385 390 395	
TAT CCC AGC GAC ATC GCC GTG GAG TGG GAG AGC AAT GGG CAG CCG GAG	1307
Tyr Pro Ser Asp Ile Ala Val Glu Trp Glu Ser Asn Gly Gln Pro Glu	
400 405 410 415	
AAC AAC TAC AAG ACC ACG CCT CCC GTG CTG GAC TCC GAC GGC TCC TTC	1355
Asn Asn Tyr Lys Thr Thr Pro Pro Val Leu Asp Ser Asp Gly Ser Phe	
420 425 430	

TTC CTC TAC AGC AAG CTC ACC GTG GAC AAG AGC AGG TGG CAG CAG GGG	1403
Phe Leu Tyr Ser Lys Leu Thr Val Asp Lys Ser Arg Trp Gln Gln Gly	
435 440 445	
AAC GTC TTC TCA TGC TCC GTG ATG CAT GAG GCT CTG CAC AAC CAC TAC	1451
Asn Val Phe Ser Cys Ser Val Met His Glu Ala Leu His Asn His Tyr	
450 455 460	
ACG CAG AAG AGC CTA AGC TTG TCT GCG GGT AAA CCC ACC CAT GTC AAT	1499
Thr Gln Lys Ser Leu Ser Leu Ser Ala Gly Lys Pro Thr His Val Asn	
465 470 475	
GTG TCT GTT GTC ATG GCG GAG GTG GAC GGC ACC TGC TAC TGATAGTCTA	1548
Val Ser Val Val Met Ala Glu Val Asp Gly Thr Cys Tyr	
480 485 490	
GAGCTCGCTG AT	1560

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 492 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Met Gly His Thr Arg Arg Gln Gly Thr Ser Pro Ser Lys Cys Pro Tyr	
1 5 10 15	
Leu Asn Phe Phe Gln Leu Leu Val Leu Ala Gly Leu Ser His Phe Cys	
20 25 30	
Ser Gly Val Ile His Val Thr Lys Glu Val Lys Glu Val Ala Thr Leu	
35 40 45	
Ser Cys Gly His Asn Val Ser Val Glu Glu Leu Ala Gln Thr Arg Ile	
50 55 60	
Tyr Trp Gln Lys Glu Lys Lys Met Val Leu Thr Met Met Ser Gly Asp	
65 70 75 80	

Met Asn Ile Trp Pro Glu Tyr Lys Asn Arg Thr Ile Phe Asp Ile Thr
 85 90 95

Asn Asn Leu Ser Ile Val Ile Leu Ala Leu Arg Pro Ser Asp Glu Gly
 100 105 110

Thr Tyr Glu Cys Val Val Leu Lys Tyr Glu Lys Asp Ala Phe Lys Arg
 115 120 125

Glu His Leu Ala Glu Val Thr Leu Ser Val Lys Ala Asp Phe Pro Thr
 130 135 140

Pro Ser Ile Ser Asp Phe Glu Ile Pro Thr Ser Asn Ile Arg Arg Ile
 145 150 155 160

Ile Cys Ser Thr Ser Gly Gly Phe Pro Glu Pro His Leu Ser Trp Leu
 165 170 175

Glu Asn Gly Glu Glu Leu Asn Ala Ile Asn Thr Thr Val Ser Gln Asp
 180 185 190

Pro Glu Thr Glu Leu Tyr Ala Val Ser Ser Lys Leu Asp Phe Asn Met
 195 200 205

Thr Thr Asn His Ser Phe Met Cys Leu Ile Lys Tyr Gly His Leu Arg
 210 215 220

Val Asn Gln Thr Phe Asn Trp Asn Thr Thr Lys Gln Glu His Phe Pro
 225 230 235 240

Asp Gln Glu Pro Lys Ser Ala Asp Lys Thr His Thr Cys Pro Pro Cys
 245 250 255

Pro Ala Pro Glu Leu Leu Gly Gly Pro Ser Val Phe Leu Phe Pro Pro
 260 265 270

Lys Pro Lys Asp Thr Leu Met Ile Ser Arg Thr Pro Glu Val Thr Cys
 275 280 285

Val Val Val Asp Val Ser His Glu Asp Pro Glu Val Lys Phe Asn Trp
 290 295 300

Tyr Val Asp Gly Val Glu Val His Asn Ala Lys Thr Lys Pro Arg Glu
 305 310 315 320

Glu Gln Tyr Asn Ser Thr Tyr Arg Val Val Ser Val Leu Thr Val Leu
 325 330 335
 His Gln Asp Trp Leu Asn Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn
 340 345 350
 Lys Ala Leu Pro Ala Pro Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly
 355 360 365
 Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser Arg Asp Glu
 370 375 380
 Leu Thr Lys Asn Gln Val Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr
 385 390 395 400
 Pro Ser Asp Ile Ala Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn
 405 410 415
 Asn Tyr Lys Thr Thr Pro Pro Val Leu Asp Ser Asp Gly Ser Phe Phe
 420 425 430
 Leu Tyr Ser Lys Leu Thr Val Asp Lys Ser Arg Trp Gln Gln Gly Asn
 435 440 445
 Val Phe Ser Cys Ser Val Met His Glu Ala Leu His Asn His Tyr Thr
 450 455 460
 Gln Lys Ser Leu Ser Leu Ser Ala Gly Lys Pro Thr His Val Asn Val
 465 470 475 480
 Ser Val Val Met Ala Glu Val Asp Gly Thr Cys Tyr
 485 490

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 831 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 52..831

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

TCGGTTACCT GCAGATATCA AGCTAATTCG GTACCAGCAG AAGCAGCCAA A ATG GAT	57
Met Asp	
1	
CCC CAG TGC ACT ATG GGA CTG AGT AAC ATT CTC TTT GTG ATG GCC TTC	105
Pro Gln Cys Thr Met Gly Leu Ser Asn Ile Leu Phe Val Met Ala Phe	
5 10 15	
CTG CTC TCT GGT GCT GCT CCT CTG AAG ATT CAA GCT TAT TTC AAT GAG	153
Leu Leu Ser Gly Ala Ala Pro Leu Lys Ile Gln Ala Tyr Phe Asn Glu	
20 25 30	
ACT GCA GAC CTG CCA TGC CAA TTT GCA AAC TCT CAA AAC CAA AGC CTG	201
Thr Ala Asp Leu Pro Cys Gln Phe Ala Asn Ser Gln Asn Gln Ser Leu	
35 40 45 50	
AGT GAG CTA GTA GTA TTT TGG CAG GAC CAG GAA AAC TTG GTT CTG AAT	249
Ser Glu Leu Val Val Phe Trp Gln Asp Gln Glu Asn Leu Val Leu Asn	
55 60 65	
GAG GTA TAC TTA GGC AAA GAG AAA TTT GAC AGT GTT CAT TCC AAG TAT	297
Glu Val Tyr Leu Gly Lys Glu Lys Phe Asp Ser Val His Ser Lys Tyr	
70 75 80	
ATG GGC CGC ACA AGT TTT GAT TCG GAC AGT TGG ACC CTG AGA CTT CAC	345
Met Gly Arg Thr Ser Phe Asp Ser Asp Ser Trp Thr Leu Arg Leu His	
85 90 95	
AAT CTT CAG ATC AAG GAC AAG GGC TTG TAT CAA TGT ATC ATC CAT CAC	393
Asn Leu Gln Ile Lys Asp Lys Gly Leu Tyr Gln Cys Ile Ile His His	
100 105 110	
AAA AAG CCC ACA GGA ATG ATT CGC ATC CAC CAG ATG AAT TCT GAA CTG	441
Lys Lys Pro Thr Gly Met Ile Arg Ile His Gln Met Asn Ser Glu Leu	
115 120 125 130	

TCA GTG CTT GCT AAC TTC AGT CAA CCT GAA ATA GTA CCA ATT TCT AAT	489
Ser Val Leu Ala Asn Phe Ser Gln Pro Glu Ile Val Pro Ile Ser Asn	
135 140 145	
ATA ACA GAA AAT GTG TAC ATA AAT TTG ACC TGC TCA TCT ATA CAC GGT	537
Ile Thr Glu Asn Val Tyr Ile Asn Leu Thr Cys Ser Ser Ile His Gly	
150 155 160	
TAC CCA GAA CCT AAG AAG ATG AGT GTT TTG CTA AGA ACC AAG AAT TCA	585
Tyr Pro Glu Pro Lys Lys Met Ser Val Leu Leu Arg Thr Lys Asn Ser	
165 170 175	
ACT ATC GAG TAT GAT GGT ATT ATG CAG AAA TCT CAA GAT AAT GTC ACA	633
Thr Ile Glu Tyr Asp Gly Ile Met Gln Lys Ser Gln Asp Asn Val Thr	
180 185 190	
GAA CTG TAC GAC GTT TCC ATC AGC TTG TCT GTT TCA TTC CCT GAT GTT	681
Glu Leu Tyr Asp Val Ser Ile Ser Leu Ser Val Ser Phe Pro Asp Val	
195 200 205 210	
ACG AGC AAT ATG ACC ATC TTC TGT ATT CTG GAA ACT GAC AAG ACG CGG	729
Thr Ser Asn Met Thr Ile Phe Cys Ile Leu Glu Thr Asp Lys Thr Arg	
215 220 225	
CTT TTA TCT TCA CCT TTC TCT ATA GAG CTT GAG GAC CCT CAG CCT CCC	777
Leu Leu Ser Ser Pro Phe Ser Ile Glu Leu Glu Asp Pro Gln Pro Pro	
230 235 240	
CCA GAC CAC GAG CCC AAA TCG GCC GAC AAA ACT CAC ACA TGC CCA CCG	825
Pro Asp His Glu Pro Lys Ser Ala Asp Lys Thr His Thr Cys Pro Pro	
245 250 255	
TGC CCA	831
Cys Pro	
260	

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 260 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

```

Met Asp Pro Gln Cys Thr Met Gly Leu Ser Asn Ile Leu Phe Val Met
 1              5              10              15

Ala Phe Leu Leu Ser Gly Ala Ala Pro Leu Lys Ile Gln Ala Tyr Phe
          20              25              30

Asn Glu Thr Ala Asp Leu Pro Cys Gln Phe Ala Asn Ser Gln Asn Gln
      35              40              45

Ser Leu Ser Glu Leu Val Val Phe Trp Gln Asp Gln Glu Asn Leu Val
      50              55              60

Leu Asn Glu Val Tyr Leu Gly Lys Glu Lys Phe Asp Ser Val His Ser
      65              70              75              80

Lys Tyr Met Gly Arg Thr Ser Phe Asp Ser Asp Ser Trp Thr Leu Arg
          85              90              95

Leu His Asn Leu Gln Ile Lys Asp Lys Gly Leu Tyr Gln Cys Ile Ile
      100              105              110

His His Lys Lys Pro Thr Gly Met Ile Arg Ile His Gln Met Asn Ser
      115              120              125

Glu Leu Ser Val Leu Ala Asn Phe Ser Gln Pro Glu Ile Val Pro Ile
      130              135              140

Ser Asn Ile Thr Glu Asn Val Tyr Ile Asn Leu Thr Cys Ser Ser Ile
      145              150              155              160

His Gly Tyr Pro Glu Pro Lys Lys Met Ser Val Leu Leu Arg Thr Lys
          165              170              175

Asn Ser Thr Ile Glu Tyr Asp Gly Ile Met Gln Lys Ser Gln Asp Asn
      180              185              190

Val Thr Glu Leu Tyr Asp Val Ser Ile Ser Leu Ser Val Ser Phe Pro
      195              200              205

Asp Val Thr Ser Asn Met Thr Ile Phe Cys Ile Leu Glu Thr Asp Lys
      210              215              220

```

Thr Arg Leu Leu Ser Ser Pro Phe Ser Ile Glu Leu Glu Asp Pro Gln
 225 230 235 240

Pro Pro Pro Asp His Glu Pro Lys Ser Ala Asp Lys Thr His Thr Cys
 245 250 255

Pro Pro Cys Pro
 260

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1104 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 601..1104

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

GACGTCGACG GATCGGGAGA TCGGGGATCG ATCCGTCGAC GTACGACTAG TTATTAATAG	60
TAATCAATTA CGGGGTCATT AGTTCATAGC CCATATATGG AGTTCCGCGT TACATAACTT	120
ACGGTAAATG GCCCGCCTGG CTGACCGCCC AACGACCCCC GCCCATTGAC GTCAATAATG	180
ACGTATGTTT CCATAGTAAC GCCAATAGGG ACTTTCCATT GACGTCAATG GGTGGACTAT	240
TTACGGTAAA CTGCCCCACTT GGCAGTACAT CAAGTGTATC ATATGCCAAG TACGCCCCCT	300
ATTGACGTCA ATGACGGTAA ATGGCCCGCC TGGCATTATG CCCAGTACAT GACCTTATGG	360
GACTTTCCTA CTTGGCAGTA CATCTACGTA TTAGTCATCG CTATTACCAT GGTGATGCGG	420
TTTTGGCAGT ACATCAATGG GCGTGGATAG CGGTTTGACT CACGGGGATT TCCAAGTCTC	480
CACCCCATTG ACGTCAATGG GAGTTTGTTC TGGCGACTCA CTATAGGAGT TCCAAGCTT	540

CTAGAGATCC CTCGAGATCC ATTGTGCTCT AAAGGACCTG AACACCGCTC CCATAAAGCC 600

ATG GCT TGC CTT GGA TTT CAG CGG CAC AAG GCT CAG CTG AAC CTG GCT 648
Met Ala Cys Leu Gly Phe Gln Arg His Lys Ala Gln Leu Asn Leu Ala
1 5 10 15

GCC AGG ACC TGG CCC TGC ACT CTC CTG TTT TTT CTT CTC TTC ATC CCT 696
Ala Arg Thr Trp Pro Cys Thr Leu Leu Phe Phe Leu Leu Phe Ile Pro
20 25 30

GTC TTC TGC AAA GCA ATG CAC GTG GCC CAG CCT GCT GTG GTA CTG GCC 744
Val Phe Cys Lys Ala Met His Val Ala Gln Pro Ala Val Val Leu Ala
35 40 45

AGC AGC CGA GGC ATC GCC AGC TTT GTG TGT GAG TAT GCA TCT CCA GGC 792
Ser Ser Arg Gly Ile Ala Ser Phe Val Cys Glu Tyr Ala Ser Pro Gly
50 55 60

AAA GCC ACT GAG GTC CGG GTG ACA GTG CTT CGG CAG GCT GAC AGC CAG 840
Lys Ala Thr Glu Val Arg Val Thr Val Leu Arg Gln Ala Asp Ser Gln
65 70 75 80

GTG ACT GAA GTC TGT GCG GCA ACC TAC ATG ACG GGG AAT GAG TTG ACC 888
Val Thr Glu Val Cys Ala Ala Thr Tyr Met Thr Gly Asn Glu Leu Thr
85 90 95

TTC CTA GAT GAT TCC ATC TGC ACG GGC ACC TCC AGT GGA AAT CAA GTG 936
Phe Leu Asp Asp Ser Ile Cys Thr Gly Thr Ser Ser Gly Asn Gln Val
100 105 110

AAC CTC ACT ATC CAA GGA CTG AGG GCC ATG GAC ACG GGA CTC TAC ATC 984
Asn Leu Thr Ile Gln Gly Leu Arg Ala Met Asp Thr Gly Leu Tyr Ile
115 120 125

TGC AAG GTG GAG CTC ATG TAC CCA CCG CCA TAC TAC CTG GGC ATA GGC 1032
Cys Lys Val Glu Leu Met Tyr Pro Pro Pro Tyr Tyr Leu Gly Ile Gly
130 135 140

AAC GGA ACC CAG ATT TAT GTA ATT GAT CCA GAA CCG TGC CCA GAT TCT 1080
Asn Gly Thr Gln Ile Tyr Val Ile Asp Pro Glu Pro Cys Pro Asp Ser
145 150 155 160

GAC GCT GAG CCC AAA TCG GCC GAC
 Asp Ala Glu Pro Lys Ser Ala Asp
 165

1104

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 168 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

Met Ala Cys Leu Gly Phe Gln Arg His Lys Ala Gln Leu Asn Leu Ala
 1 5 10 15

Ala Arg Thr Trp Pro Cys Thr Leu Leu Phe Phe Leu Leu Phe Ile Pro
 20 25 30

Val Phe Cys Lys Ala Met His Val Ala Gln Pro Ala Val Val Leu Ala
 35 40 45

Ser Ser Arg Gly Ile Ala Ser Phe Val Cys Glu Tyr Ala Ser Pro Gly
 50 55 60

Lys Ala Thr Glu Val Arg Val Thr Val Leu Arg Gln Ala Asp Ser Gln
 65 70 75 80

Val Thr Glu Val Cys Ala Ala Thr Tyr Met Thr Gly Asn Glu Leu Thr
 85 90 95

Phe Leu Asp Asp Ser Ile Cys Thr Gly Thr Ser Ser Gly Asn Gln Val
 100 105 110

Asn Leu Thr Ile Gln Gly Leu Arg Ala Met Asp Thr Gly Leu Tyr Ile
 115 120 125

Cys Lys Val Glu Leu Met Tyr Pro Pro Pro Tyr Tyr Leu Gly Ile Gly
 130 135 140

Asn Gly Thr Gln Ile Tyr Val Ile Asp Pro Glu Pro Cys Pro Asp Ser
145 150 155 160

Asp Ala Glu Pro Lys Ser Ala Asp
165

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Pro Thr His Val Asn Val Ser Val Val Met Ala Glu Val Asp Gly Thr
1 5 10 15

Cys Tyr

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 6 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

Ser Leu Ser Pro Gly Lys
1 5

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single

(D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Pro	Thr	Leu	Tyr	Asn	Val	Ser	Leu	Val	Met	Ser	Asp	Thr	Ala	Gly	Thr
1				5					10					15	

Cys Tyr

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 18 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

Pro	Thr	His	Val	Asn	Val	Ser	Val	Val	Met	Ala	Glu	Val	Asp	Gly	Thr
1				5					10					15	

Cys Tyr

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 10 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

Gly	Pro	Ser	Lys	Pro	Glu	Pro	Lys	Ser	Ala
1				5					10

(2) INFORMATION FOR SEQ ID NO:13:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 4 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

Asn Lys Ile Leu

1

(2) INFORMATION FOR SEQ ID NO:14:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 10 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

His Phe Pro Asp Gln Glu Pro Lys Ser Ala

1

5

10

(2) INFORMATION FOR SEQ ID NO:15:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 4 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

Val Ile His Val

1

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

Pro	Pro	Pro	Asp	His	Glu	Pro	Lys	Ser	Ala
1				5					10

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 4 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

Leu	Lys	Ile	Gln
1			

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 13 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

Glu	Pro	Cys	Pro	Asp	Ser	Asp	Ala	Glu	Pro	Lys	Ser	Ala
1				5								10

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 4 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

Met His Val Ala

1

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

CCCAAATCGG CCGACAAAAC T

21

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 57 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

TCAGCGAGCT CTAGACTACA CTCATTACC CGGAGACAAG CTTAGGCTCT TCTGCGT

57

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 79 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

AGCTTGCTCTG CGGGTAAACC CACCCATGTC AATGTGCTCTG TTGTCATGGC GGAGGTGGAC 60
GGCACCTGCT ACTGATAGT 79

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 79 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

CTAGACTATC AGTAGCAGGT GCCGTCCACC TCCGCCATGA CAACAGACAC ATTGACATGG 60
GTGGGTTTAC CCGCAGACA 79

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 14 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

Gly Pro Ser Lys Pro Glu Pro Lys Ser Ala Gly Ile Lys Pro
1 5 10

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 6 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

Ser Leu Ser Thr Gly Lys
1 5

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 6 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

Ser Leu Ser Ala Gly Lys
1 5

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 5 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

Glu Pro Lys Ser Ala
1 5

What is claimed is:

1. A hexameric fusion protein comprising:
 - (a) a dimeric binding protein and
 - 5 (b) a tailpiece α p) characterized by having the activity of the tailpiece from the C-terminus of the heavy chain of an IgA antibody.
2. The fusion protein according to claim 1, wherein the dimeric binding protein is selected from the group consisting of:
 - 10 (a) a protein fragment comprising the extracellular domain of a selected monomeric binding protein or a functional fragment thereof fused to an Ig-Fc fragment selected from the group consisting of an Fc fragment from an IgG antibody, an Fc fragment from an IgD antibody, an Fc fragment from an IgE antibody, and an Fc fragment from an IgM antibody excluding the μ tp; and
 - 15 (b) a naturally dimeric binding protein or a fragment thereof having the binding ability of said dimeric protein.
3. The fusion protein according to claim 2 further comprising a leader suitable for expression and processing of the fusion protein.
- 20 4. The fusion protein according to claim 2 wherein the protein fragment consists of the native leader and extracellular domains selected from the group consisting of CD80, CTLA-4 and CD86.
- 25 5. The fusion protein according to claim 2 wherein the dimeric binding protein is a Ig-Fab fragment and the heavy chain is joined to the Ig-Fc fragment.
6. The fusion protein according to claim 2 wherein the Ig-Fc fragment is from an IgG antibody selected from the group of human isotypes consisting of IgG₁, IgG₂, IgG₃,
30 IgG₄, and IgG binding mutants.
7. The fusion protein according to claim 2 wherein the Ig-Fc fragment is from a human IgG₁ antibody.

8. The fusion protein according to claim 1 wherein the α tp is the tailpiece of an antibody selected from the group consisting of human IgA1, human IgA2, rabbit IgA, mouse IgA, and gorilla IgG.
- 5 9. The fusion protein according to claim 8 wherein the α tp has the sequence SEQ ID NO: 10 PTHVNVSVVMAEVDGTCY.
- 10 10. The fusion protein according to claim 8 wherein the α tp has been modified to remove the N-linked glycosylation site.
11. The fusion protein according to claim 1 further comprising a linker of between 1 to about 20 amino acids in length, said linker located between the binding protein and the α tp.
- 15 12. The fusion protein according to claim 1 which is a homo-hexamer.
13. The fusion protein according to claim 1 which is a hetero-hexamer.
- 20 14. A polynucleotide sequence encoding a hexameric fusion protein comprising:
(a) a dimeric binding protein and
(b) a tailpiece (α tp) characterized by having the biological activity of the tailpiece from the C-terminus of the heavy chain of an IgA antibody.
- 25 15. The polynucleotide sequence according to claim 14, wherein the dimeric binding protein is selected from the group consisting of:
(a) a protein fragment comprising the extracellular domain of a selected monomeric binding protein fused to an Ig-Fc fragment selected from the group consisting of an Fc fragment from an IgG antibody, an Fc fragment from an IgD antibody, an Fc fragment from an IgE antibody, and an Fc fragment from an IgM antibody excluding the μ tp; and
30 (b) a naturally dimeric binding protein or a fragment thereof having the binding ability of said protein.

16. The polynucleotide sequence according to claim 15 further comprising a leader suitable for expression and processing of the fusion protein.
17. The polynucleotide sequence according to claim 15 wherein the protein
5 fragment consists of the native leader and extracellular domains selected from the group consisting of CD80, CTLA-4 and CD86.
18. The polynucleotide sequence according to claim 15 wherein the dimeric
10 protein is a Ig-Fab fragment and the heavy chain is joined to the Ig-Fc fragment.
19. The polynucleotide sequence according to claim 15 wherein the Ig-Fc
fragment is from an IgG antibody selected from the group of human isotypes consisting of
IgG₁, IgG₂, IgG₃, IgG₄, and IgG binding mutants.
20. The polynucleotide sequence according to claim 15 wherein the Ig-Fc
15 fragment is from a human IgG₁ antibody.
21. The polynucleotide sequence according to claim 14 wherein the α tp is the
tailpiece of an antibody selected from the group consisting of human IgA1, human IgA2,
20 rabbit IgA, mouse IgA, and gorilla IgG.
22. The polynucleotide sequence according to claim 21 wherein the α tp has the
sequence SEQ ID NO: 101 THVNVSVVMAEVDGTCY.
23. The polynucleotide sequence according to claim 21 wherein the α tp has been
25 modified to remove the N-linked glycosylation site.
24. The polynucleotide sequence according to claim 14, wherein the fusion
protein further comprises a linker of between 1 to about 20 amino acids in length, said linker
30 located between the binding protein and the α tp.
25. A vector comprising a polynucleotide sequence encoding:
(a) a polynucleotide sequence according to claim 14; and

(b) sequences controlling expression of the fusion protein in a selected host cell.

5

26. A recombinant host cell comprising the vector of claim 25.

27. A pharmaceutical composition comprising an hexameric fusion protein according to claim 1 in a pharmaceutically acceptable carrier.

10 28. A pharmaceutical composition comprising a polynucleotide sequence according to claim 14 in a pharmaceutically acceptable carrier.

29. A diagnostic reagent comprising a detectable label and an hexameric fusion protein according to claim 1.

15 30. A diagnostic reagent comprising a detectable label and a polynucleotide sequence according to claim 14.

20 31. A method for producing a hexameric fusion protein comprising the steps of:
(a) providing a dimeric binding protein; and
(b) attaching to each monomer of said binding protein a tailpiece (otp) characterized by having the biological activity of the tailpiece from the C-terminus of the heavy chain of an IgA antibody.

25 32. A method of purifying a hexameric fusion protein comprising:
(a) providing a selected host cell according to claim 26;
(b) recovering the stable hexameric fusion protein; and
(c) purifying the recovered fusion protein.

30 33. The method according to claim 32, wherein said fusion protein comprises IgG or a fragment thereof, and said purification step comprises the step of applying said fusion protein to a Protein A or Protein G column.

34. A method for screening for a ligand which binds to a hexameric fusion protein according to claim 1, comprising the steps of:
- (a) providing the hexameric fusion protein;
 - (b) permitting a test sample to come into contact with the hexameric fusion protein; and
 - (c) detecting binding between the fusion protein and any ligand in the test sample.
35. The method according to claim 34 wherein the fusion protein is immobilized to a surface.
36. The method according to claim 34 wherein the fusion protein is in solution.
37. The method according to claim 34, wherein the fusion protein is selected from the group consisting of CD80-Igαtp and CD86-Igαtp.
38. A method for screening for a compound that inhibits the interaction between a selected binding protein and a ligand, said method comprising the step of
- (a) providing a known ligand for said binding protein;
 - (b) providing a hexameric fusion protein according to claim 1;
 - (c) contacting the known ligand with a test solution;
 - (d) contacting the known ligand with the hexameric fusion protein;
 - (e) detecting inhibition of binding of the hexameric fusion protein; and
 - (f) optionally isolating the compound which binds to the hexameric protein.
39. The method according to claim 38, wherein the ligand is selected from the group consisting of CD28 and CTLA-4 and the hexameric fusion protein is selected from the group consisting of CD80-Igαtp and CD86-Igαtp.
40. A method for stimulating CD28 positive cells comprising the step of administering to CD28 positive cells a hexameric fusion protein selected from the group consisting of CD80-Igαtp and CD86-Igαtp.

41. A method for suppressing CTLA-4 positive cells comprising the step of administering to CTLA-4 positive cells a hexameric fusion protein selected from the group consisting of CD80-Ig α p and CD86-Ig α p.
- 5
42. A method for antagonizing cell surface CD80- and CD86-mediated stimulation of CD28 positive cells by administering to said cells a hexameric fusion protein CTLA4-Ig α p.
- 10
43. A method for immunizing an animal comprising the method of administering to the animal an effective amount of a pharmaceutical compositions according to claim 27.
44. A method for immunizing an animal comprising the method of administering to the animal an effective amount of a pharmaceutical compositions according to claim 28.

FIG. 1

Hexameric CD80-Ig α tp

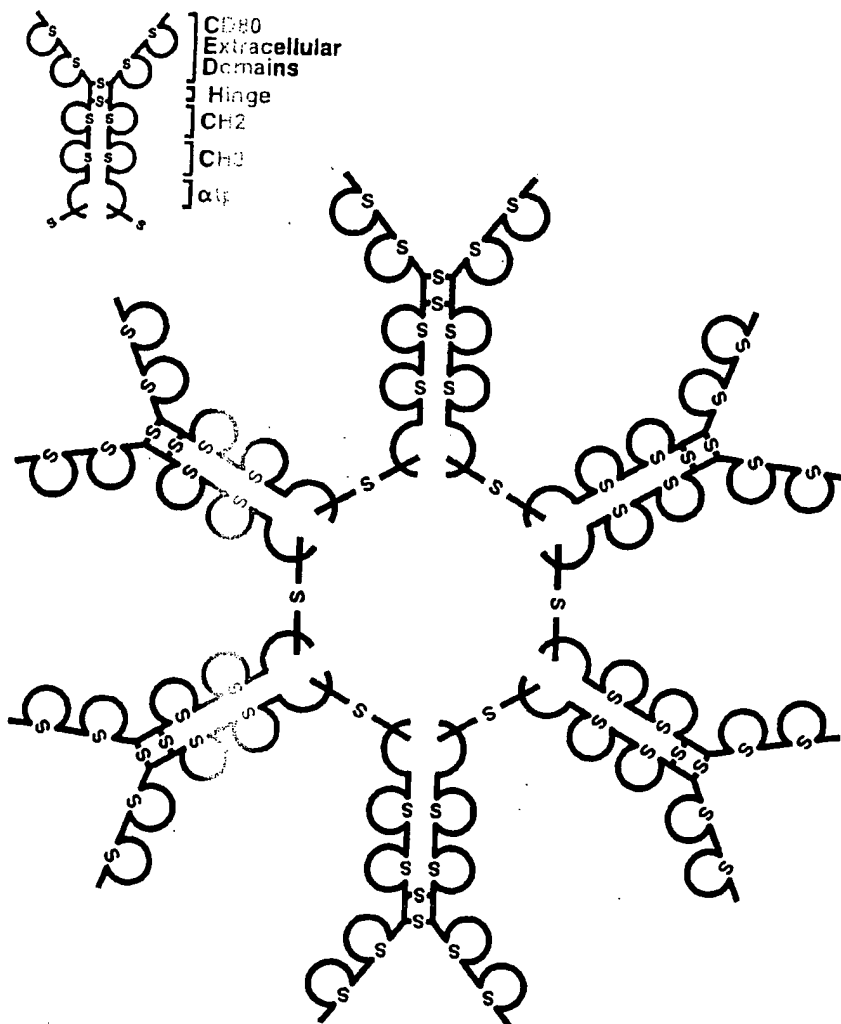
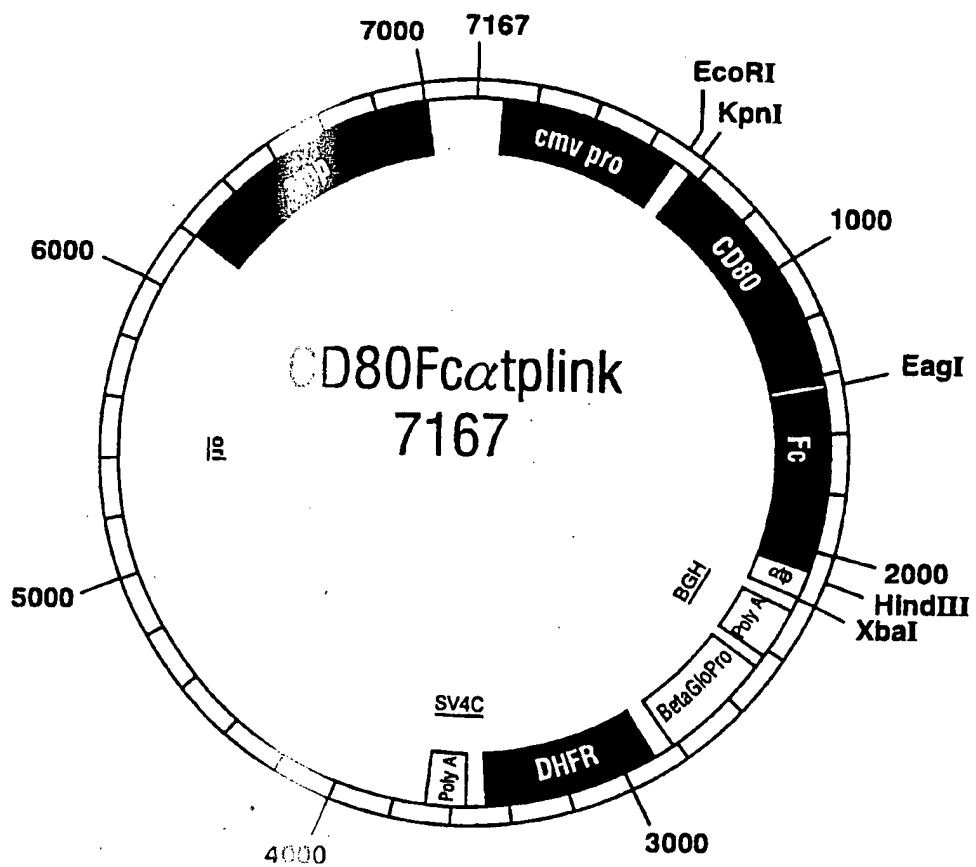


FIG. 2



3/35

FIG. 3

DNA sequence of CD8 α -Scatplink

1 GACGTCGAC GATCGGGAGA TCGGGGATCG ATCCGTCGAC GTACGACTAG
51 TTATTAATAA TAATCAATTA CGGGGTCATT AGTTCATAGC CCATATATGG
101 AGTTCCGCC TACATAACTT ACGGTAAATG GCCCGCCTGG CTGACCGCCC
151 AACGACCCC GCCCATTGAC GTCAATAATG ACGTATGTTC CCATAGTAAC
201 GCCAATAGGC ACTTTCCATT GACGTCAATG GGTGGACTAT TTACGGTAAA
251 CTGCCCCACTT GGCAGTACAT CAAGTGTATC ATATGCCAAG TACGCCCCCT
301 ATTGACGTG ATGACGGTAA ATGGCCCGCC TGGCATTATG CCCAGTACAT
351 GACCTTATC GACTTTCCTA CTTGGCAGTA CATCTACGTA TTAGTCATCG
401 CTATTACCAG GGTGATGCGG TTTTGGCAGT ACATCAATGG GCGTGGATAG
451 CGGTTTGACT CACGGGGATT TCCAAGTCTC CACCCCATTG ACGTCAATGG
501 GAGTTTGTTT TGGCACCAAA ATCAACGGGA CTTTCCAAA TGTGTAACA
551 ACTCCGCCG ATTGACGCAA ATGGGCGGTA GGCGTGTACG GTGGGAGGTC
601 TATATAAGCA GAGCTGGGTA CGTGAACCGT CAGATCGCCT GGAGACGCCA
651 TCGAATTGG TTACCTGCAG ATATCAAGCT AATTCGGTAC CGGGCCCCC
701 TCGAGCCTCA AGCCATGGGC CACACACGGA GGCAGGGAAC ATCACCATCC
751 AAGTGTCAT ACCTCAATTT CTTTCAGCTC TTGGTGCTGG CTGGTCTTTC
801 TCACTTCTGT TCAGGTGTTA TCCACGTGAC CAAGGAAGTG AAAGAAGTGG
851 CAACGCTGTC CTGTGGTCAC AATGTTTCTG TTGAAGAGCT GGCACAACT
901 CGCATCTACT GCAAAAAGGA GAAGAAAATG GTGCTGACTA TGATGTCTGG

FIG. 3A

951 GGACATGAAT ATATGGCCCG AGTACAAGAA CCGGACCATC TTTGATATCA
1001 CTAATAACCT CTCCATTGTG ATCCTGGCTC TGCGCCCATC TGACGAGGGC
1051 ACATACGAA GTGTTGTTCT GAAGTATGAA AAAGACGCTT TCAAGCGGGA
1101 ACACCTGGC GAAGTGACGT TATCAGTCAA AGCTGACTTC CCTACACCTA
1151 GTATATCTG CTTTGAAATT CCAACTTCTA ATATTAGAAG GATAATTTGC
1201 TCAACCTCT GAGGTTTTCC AGAGCCTCAC CTCTCCTGGT TGGAAAATGG
1251 AGAAGAATT AATGCCATCA ACACAACAGT TTCCCAAGAT CCTGAAACTG
1301 AGCTCTATG TGTAGCAGC AACTGGATT TCAATATGAC AACCAACCAC
1351 AGCTTCATC GTCTCATCAA GTATGGACAT TTAAGAGTGA ATCAGACCTT
1401 CAACTGGAA ACAACCAAGC AAGAGCATT TCCCTGATCAG GAGCCCAAAT
1451 CGGCCGACA AACTCACACA TGCCACCGT GCCCAGCACC TGAATCCTG
1501 GGGGGACCG CAGTCTTCCT CTCCCCCA AAACCAAGG ACACCCTCAT
1551 GATCTCCCG ACCCCTGAGG TCACATGCGT GGTGGTGGAC GTGAGCCACG
1601 AAGACCCTG GGTCAAGTTC AACTGGTACG TGGACGGCGT GGAGGTGCAT
1651 AATGCCAAGA CAAAGCCGCG GGAGGAGCAG TACAACAGCA CGTACCGGGT
1701 GGTCAGCGT CTCACCGTCC TGCACCAGGA CTGGCTGAAT GGCAAGGAGT
1751 ACAAGTGCA GGTCTCCAAC AAAGCCCTCC CAGCCCCCAT CGAGAAAACC
1801 ATCTCCAAAG TCAAAGGGCA GCCCCGAGAA CCACAGGTGT ACACCCTGCC
1851 CCCATCCCG CATGAGCTGA CCAAGAACCA GGTCAGCCTG ACCTGCCTGG
1901 TCAAAGGCT TATCCCAGC GACATCGCG TGGAGTGGGA GAGCAATGGG

FIG. 3B

1951 CAGCCGGAA ACAACTACAA GACCACGCCT CCCGTGCTGG ACTCCGACGG
2001 CTCCTTCTT CTCTACAGCA AGCTCACCGT GGACAAGAGC AGGTGGCAGC
2051 AGGGGAACCTT CTCTCATGC TCCGTGATGC ATGAGGCTCT GCACAACCAC
2101 TACACGCATA AGAGCCTAAG CTTGTCTGCG GGTAAACCCA CCCATGTCAA
2151 TGTGTCTGT TGCATGGCGG AGGTGGACGG CACCTGCTAC TGATAGTCTA
2201 GAGCTGCGTA ATCAGCCTCG ACTGTGCCTT CTAGTTGCCA GCCATCTGTT
2251 GTTTGCCCCCT CCCCCGTGCC TTCCTTGACC CTGGAAGGTG CCACTCCCAC
2301 TGTCCCTTT TAATAAAATG AGGAAATTGC ATCGCATTGT CTGAGTAGGT
2351 GTCATTCTAT TCTGGGGGGT GGGGTGGGGC AGGACAGCAA GGGGGAGGAT
2401 TGGGAAGACA ATAGCAGGCA TGCTGGGGAT GCGGTGGGCT CTATGGAACC
2451 AGCTGGGGCT CGAGGGGGGA TCTCCCGATC CCCAGCTTTG CTTCTCAATT
2501 TCTTATTTCC ATAATGAGAA AAAAAGGAAA ATTAATTTTA ACACCAATTC
2551 AGTAGTTGA TGAGCAAATG CGTTGCCAAA AAGGATGCTT TAGAGACAGT
2601 GTTCTCTGCA CAGATAAGGA CAAACATTAT TCAGAGGGAG TACCCAGAGC
2651 TGAGACTCC AAGCCAGTGA GTGGCACAGC ATTCTAGGGA GAAATATGCT
2701 TGTCAACAG GAAGCCTGAT TCCGTAGAGC CACACCTTGG TAAGGGCCAA
2751 TCTGCTCACA CAGGATAGAG AGGGCAGGAG CCAGGGCAGA GCATATAAGG
2801 TGAGGTAGC TCAGTTGCTC CTCACATTG CTTCTGACAT AGTTGTGTTG
2851 GGAGCTTGCA TAGCTTGAC AGCTCAGGGC TGCGATTTG CGCCAACTT
2901 GACGGCAAT CTAGCGTGAA GGCTGGTAGG ATTTTATCCC CGCTGCCATC

FIG. 3C

2951 ATGGTTCGAC CATTGAACTG CATCGTCGCC GTGTCCCAA ATATGGGGAT
3001 TGGCAAGAC GAGACCTAC CCTGGCCTCC GCTCAGGAAC GAGTTCAAGT
3051 ACTTCCAA AATGACCACA ACCTCTTCAG TGGAAGGTAA ACAGAATCTG
3101 GTGATTATAG TAGGAAAAC CTGGTTCTCC ATTCCTGAGA AGAATCGACC
3151 TTTAAAGGAG AGAATTAATA TAGTTCTCAG TAGAGAACTC AAAGAACCAC
3201 CACGAGGAT TCATTTTCTT GCCAAAAGTT TGGATGATGC CTTAAGACTT
3251 ATTGAACAA CGGAATTGGC AAGTAAAGTA GACATGGTTT GGATAGTCGG
3301 AGGCAGTTCT OTTTACCAGG AAGCCATGAA TCAACCAGGC CACCTTAGAC
3351 TCTTTGTGAG AAGGATCATG CAGGAATTG AAAGTGACAC GTTTTCCCA
3401 GAAATTGAT TGGGGAAATA TAACTTCTC CCAGAATACC CAGGCGTCTT
3451 CTCTGAGGT CAGGAGGAAA AAGGCATCAA GTATAAGTTT GAAGTCTACG
3501 AGAAGAAAG CTAACAGGAA GATGCTTTCA AGTTCTCTGC TCCCCTCCTA
3551 AAGCTATGC TTTTATAAG ACCATGCTAG CTTGAACTTG TTTATTGCAG
3601 CTTATAATG CTACAAATAA AGCAATAGCA TCACAAATTT CACAAATAAA
3651 GCATTTTTC TACTGCATTC TAGTTGTGGT TTGTCCAAAC TCATCAATGT
3701 ATCTTATCAT CTCTGGATCA ACGATAGCTT ATCTGTGGGC GATGCCAAGC
3751 ACCTGGATG GTTGGTTTC CTGCTACTGA TTTAGAAGCC ATTTGCCCCC
3801 TGAGTGGGCA TGGGAGCAC TAACTTTCTC TTCAAAGGA AGCAATGCAG
3851 AAAGAAAAG ATACAAAGTA TAAGCTGCCA TGTAATAATG GAAGAAGATA
3901 AGGTTGTAT AATTAGATT ACATACTTCT GAATTGAAAC TAAACACCTT

7/35

FIG. 3D

3951 TAAATTCCT AATATATAAC ACATTTTCATA TGAAAGTATT TTACATAAGT
4001 AACTCAGAT CATAGAAAAC AAAGCTAATG ATAGGTGTCC CTAAAAGTTC
4051 ATTTATTAT TCTACAAATG ATGAGCTGGC CATCAAAATT CCAGCTCAAT
4101 TCTTCAACCA ATTAGAAAGA GCAATCTGCA AACTCATCTG GAATAACAAA
4151 AAACCTAGAT TAGCAAAAAC TCTTCTCAAG GATAAAAGAA CCTCTGGTGG
4201 AATCACCAAT CCTGACCTAA AGCTGTACTA CAGAGCAATT GTGATAAAAA
4251 CTGCATGGTT CTGATATAGA AACGGACAAG TAGACCAATG GAATAGAACC
4301 CACACACCTT TGGTCACTTG ATCTTCAACA AGAGAGCTAA AACCATCCAC
4351 TGGAAAAAT ACAGCATTTT CAACAAATGG TGCTGGCACA ACTGGTGGTT
4401 ATCATGGAT AGAATGTGAA TTGATCCATT CCAATCTCCT TGTACTAAGG
4451 TCAAATCTAT TTGGATCAAG GAACTCCACA TAAAACCAGA GAACTGAAA
4501 CTTATAGAGT AGAAAGTGGG GAAAAGCCTC GAAGATATGG GCACAGGGGA
4551 AAAATTCCT AATAGAACAG CAATGGCTTG TGCTGTAAGA TCGAGAATTG
4601 ACAATGGG CCTCATGAAA CTCCAAAGCT ATCGGATCAA TTCCTCCAA
4651 AAAGCCTCCT CACTACTTCT GGAATAGCTC AGAGGCCGAG GCGGCCTCGG
4701 CCTCTGCATA AATAAAAAAA ATTAGTCAGC CATGCATGGG GCGGAGAATG
4751 GCGGGAACCT GCGGAGTTA GGGGCGGGAT GGGCGGAGTT AGGGGCGGGA
4801 CTATGGTTC TGACTAATTG AGATGCATGC TTTGCATACT TCTGCCTGCT
4851 GGGGAGCCTT GGGACTTTCC ACACCTGGTT GCTGACTAAT TGAGATGCAT
4901 GCTTTGCATA GTTCTGCCTG CTGGGGAGCC TGGGGACTTT CCACACCCTA

8/35

FIG. 3E

4951 ACTGACACAG ATTCCACAGA ATTAATTCCC GATCCCGTCG ACCTCGAGAG
5001 CTTGGCGTA TCATGGTCAT AGCTGTTTCC TGTGTGAAAT TGTTATCCGC
5051 TCACAATTC ACACAACATA CGAGCCGGAA GCATAAAGTG TAAAGCCTGG
5101 GGTGCCTAAT GAGTGAGCTA ACTCACATTA ATTGCGTTGC GCTCACTGCC
5151 CGCTTTCCAT TCGGGAAACC TGTCGTGCCA GCTGCATTAA TGAATCGGCC
5201 AACGCGCGG AGAGGCGGT TTGCGTATTG GCGCTCTTC CGCTTCCTCG
5251 CTCACTGAC TCGCTGCGCTC GGTGCTTCGG CTGCGGCGAG CGGTATCAGC
5301 TCACTCAAA TCGGTAATAC GGTATCCAC AGAATCAGGG GATAACGCAG
5351 GAAAGAACA CTGAGCAAAA GGCCAGCAAA AGGCCAGGAA CCGTAAAAAG
5401 GCCGCGTTG TGGCGTTTTT CCATAGGCTC CGCCCCCTG ACGAGCATCA
5451 CAAAATCGT TCGCTCAAGTC AGAGGTGGCG AAACCCGACA GGACTATAAA
5501 GATACCAGC TTTTCCCCCT GGAAGCTCCC TCGTGCCTC TCCTGTTCCG
5551 ACCCTGCCG TTACCGGATA CCTGTCCGCC TTTCTCCCTT CGGGAAGCGT
5601 GCGCTTTC TAATGCTCAC GCTGTAGGTA TCTCAGTTCG GTGTAGGTCTG
5651 TTCGCTCCAT TGTGGGCTGT GTGCACGAAC CCCCCGTTCA GCCCCACCGC
5701 TGCGCCTTA TGGGTAAC TA TCGTCTTGAG TCCAACCCGG TAAGACACGA
5751 CTTATCGCCA CTGGCAGCAG CCACTGGTAA CAGGATTAGC AGAGCGAGGT
5801 ATGTAGGCG TCGTACAGAG TTCTGAAGT GGTGGCCTAA CTACGGCTAC
5851 ACTAGAAGC TAGTATTTGG TATCTGCGCT CTGCTGAAGC CAGTTACCTT
5901 CGGAAAAGC TTGGTAGCT CTTGATCCGG CAAACAAACC ACCGCTGGTA

9/35

FIG. 3F

5951 GCGGTGGT TTTTGGTTGC AAGCAGCAGA TTACGCGCAG AAAAAAAGGA
6001 TCTCAAGA ATCCTTTGAT CTTTCTACG GGGTCTGACG CTCAGTGGAA
6051 CGAAACT GTTAAGGGA TTTTGGTCAT GAGATTATCA AAAAGGATCT
6101 TCACCTAGT CTTTAAAT TAAAAATGAA GTTTAAATC AATCTAAAGT
6151 ATATATGA AACTTGGTC TGACAGTTAC CAATGCTTAA TCAGTGAGGC
6201 ACCTATCTT GCGATCTGTC TATTCGTTC ATCCATAGTT GCCTGACTCC
6251 CCGTCGTGT CATAACTACG ATACGGGAGG GCTTACCATC TGGCCCCAGT
6301 GCTGCAATG TACCGCGAGA CCCACGCTCA CCGGCTCCAG ATTTATCAGC
6351 AATAAACCG CAGCCGGAA GGGCCGAGCG CAGAAGTGGT CCTGCAACTT
6401 TATCCGCC ATCCAGTCT ATTAATTGTT GCCGGAAGC TAGAGTAACT
6451 AGTCGCCA TTAATAGTTT GCGCAACGTT GTTGCCATTG CTACAGGCAT
6501 CGTGGTGT GCTCGTCGT TTGGTATGGC TTCATTACG TCCGGTTCCC
6551 AACGATCA GCGAGTTACA TGATCCCCCA TGTGTGCAA AAAAGCGGTT
6601 AGCTCCTT CTCTCCGAT CGTTGTCAGA AGTAAGTTGG CCGCAGTGTT
6651 ATCACTCAT GTTATGGCAG CACTGCATAA TTCTCTTACT GTCATGCCAT
6701 CCGTAAGAT CTTTCTGTG ACTGGTGAGT ACTCAACCAA GTCATTCTGA
6751 GAATAGTGT GCGGCGACC GAGTTGCTCT TGCCCGGCGT CAATACGGGA
6801 TAATACCGG ACATAGCA GAACTTTAAA AGTGCTCATC ATTGGAAAAC
6851 GTTCTTCGG GAAAACTC TCAAGGATCT TACCGCTGTT GAGATCCAGT
6901 TCGATGTAAC TACTCGTGC ACCCAACTGA TCTTCAGCAT CTTTACTTT

10/35

FIG. 3G

6951 CACCAGCGTT TCTGGGTGAG CAAAAACAGG AAGGCAAAAT GCCGCAAAAA
7001 AGGGAATAAG GCGGACACGG AAATGTTGAA TACTCATACT CTCCTTTTT
7051 CAATATTATT AAGCATTTA TCAGGGTTAT TGTCTCATGA GCGGATACAT
7101 ATTTGAATGT ATTTAGAAAA ATAAACAAAT AGGGGTCCG CGCACATTTC
7151 CCCGAAAAGT CCACCT

FIG. 4

Kpn I

Eco RI
GAATTCGGTTACCGCAGATATCAAGCTAATTCCGGTACCGGGCCCCCCCCTCGAGCCTGAAG
-----+-----+-----+-----+-----+-----712
CTTAAGCCAATGACGCTCTATAGTTCGATTAAGCCATGGCCCCGGGGGGAGCTCGGACTTC

CCATGGGGCCACAACGGAGGCAGGGAACATCACCATCCAAGTGTCATACCTCAATTTCT
-----+-----+-----+-----+-----+-----772
GGTACCCGGGTGGCGCCTCCGTCCCTTGTAGTGGTAGGTTACAGGTATGGAGTTAAAGA
Initiation codon
M G H T R R Q G T S P S K C P Y L N F F -

TTCAGCTCTTGGCTCTGGCTGGTCTTTCTCACTTCTGTTCAGGTGTTATCCACGTGACCA
-----+-----+-----+-----+-----+-----832
AAGTCGAGAACCAGACCGACCAGAAAGAGTGAAGACAAGTCCACAATAGGTGCACTGGT
Mature Peptide Seq.
Q L L V L A G L S H F C S G V I H V T K -

AGGAAGTGAAAGAGTGGCAACGCTGTCCTGTGGTCACAATGTTTCTGTTGAAGAGCTGG
-----+-----+-----+-----+-----+-----892
TCCTTCACTTTCTCACCGTTGCGACAGGACACCAGTGTTACAAAGACAACCTTCTCGACC
E V K E V A T L S C G H N V S V E E L A -

CACAAACTCGCATACTGGCAAAGGAGAAGAAAATGGTGCTGACTATGATGTCTGGGG
-----+-----+-----+-----+-----+-----952
GTGTTTGAGCGTATGACCGTTTTCTCTTCTTTTACCACGACTGATACTACAGACCCC
Q T R I Y W Q K E K K M V L T M M S G D -

ACATGAATATATCCCCGAGTACAAGAACCGGACCATCTTTGATATCACTAATAACCTCT
-----+-----+-----+-----+-----+-----1012
GTACTTATATAGGGCTCATGTCTTGGCCTGGTAGAACTATAGTGATTATTGGAGA
M N I W P E Y K N R T I F D I T N N L S -

CATGTGATCCCGCTCTGCGCCCATCTGACGAGGGCACATACGAGTGTGTTGTTCTGA
-----+-----+-----+-----+-----+-----1072
GTAACACTAGGACGAGACGCGGGTAGACTGCTCCCGTGTATGCTCACACAACAAGACT
I V I L A L R P S D E G T Y E C V V L K -

12/35

FIG. 4A

```

AGTATGAAAAACCGCTTTCAAGCGGGAACACCTGGCTGAAGTGACGTTATCAGTCAAAG
1073 -----+-----+-----+-----+-----+-----+----- 1132
TCATACTTTTTCGCGAAAGTTCGCCCTTGTGGACCGACTTCACTGCAATAGTCAGTTTC
Y E K I A F K R E H L A E V T L S V K A -

CTGACTTCCCTAACCTAGTATATCTGACTTTGAAATTCCTAATATTAGAAGGA
1133 -----+-----+-----+-----+-----+-----+----- 1192
GACTGAAGGGATGGATCATATAGACTGAAACTTTAAGGTTGAAGATTATAATCTTCCT
D F P T P S I S D F E I P T S N I R R I -

TAATTTGCTCAATCTCTGGAGGTTTCCAGAGCCTCACCTCTCCTGGTTGAAAAATGGAG
1193 -----+-----+-----+-----+-----+-----+----- 1252
ATTAAACGAGTTGAGACCTCCAAAAGGTCTCGGAGTGGAGAGGACCAACCTTTTACCTC
I C S T S G G F P E P H L S W L E N G E -

AAGAATTAAATGATCAACACAACAGTTTCCCAAGATCCTGAAACTGAGCTCTATGCTG
1253 -----+-----+-----+-----+-----+-----+----- 1312
TTCTTAATTTACATAGTTGTGTTGTCAAAGGGTTCTAGGACTTTGACTCGAGATACGAC
E L N A I N T T V S Q D P E T E L Y A V -

TTAGCAGCAAACGATTTCATATGACAACCAACCACAGCTTCATGTGTCTCATCAAGT
1313 -----+-----+-----+-----+-----+-----+----- 1372
AATCGTCGTTGCTAAAGTTATACTGTTGGTTGGTGTGCGAAGTACACAGAGTAGTTCA
S S K L D F N M T T N H S F M C L I K Y -

ATGGACATTTAATGTGAATCAGACCTTCAACTGGAATACAACCAAGCAAGAGCATTTTC
1373 -----+-----+-----+-----+-----+-----+----- 1432
TACCTGTAAATTCACTTAGTCTGGAAGTTGACCTTATGTTGGTTCGTTCTCGTAAAG
G H L R V N Q T F N W N T T K Q E H F P -

```

Eag I

```

CTGATCAGGAGCGTAAATCGGCCGACAAAACCTCACACATGCCCACCGTGCCCAGCACCTG
1433 -----+-----+-----+-----+-----+-----+----- 1492
GACTAGTCCTCGCTTAGCCGGCTGTTTTGAGTGTGTACGGGTGGCACGGGTCGTGGAC

```

Ig F_H gene

```

D Q E P K S A D K T H T C P P C P A P E -

```

AACTCCTGGGGGACCGTCAGTCTTCTCTCCCCCAAAACCCAAGGACACCCTCATGA
 1493 -----+-----+-----+-----+-----+-----+-----+----- 1552
 TTGAGGACCCCGGGCAGTCAGAAGGAGAAGGGGGTTTTGGGTTCCTGTGGGAGTACT
 L L G G P S V F L F P P K P K D T L M I -
 TCTCCCGGACCCCGAGGTCACATGCGTGGTGGTGGACGTGAGCCACGAAGACCCTGAGG
 1553 -----+-----+-----+-----+-----+-----+-----+----- 1612
 AGAGGGCCTGGGACTCCAGTGTACGCACCACCACCTGCACTCGGTGCTTCTGGGACTCC
 S R T P E V T C V V V D V S H E D P E V -
 TCAAGTTCAACTATACGTGGACGGCGTGGAGGTGCATAATGCCAAGACAAAGCCGCGGG
 1613 -----+-----+-----+-----+-----+-----+-----+----- 1672
 AGTTCAAGTTGACATGCACCTGCCGCACCTCCACGTATTACGGTTCTGTTTCGGCGCCC
 K F N W Y V D G V E V H N A K T K P R E -
 AGGAGCAGTACAGAGCAGTACCGGGTGGTTCAGCGTCTCACCGTCTGCACCAGGACT
 1673 -----+-----+-----+-----+-----+-----+-----+----- 1732
 TCCTCGTCATGTTTCGTGCATGGCCACCAAGTTCGAGGAGTGGCAGGACGTGGTCTCTGA
 E Q Y N S T Y R V V S V L T V L H Q D W -
 GGCTGAATGGCAGAGAGTACAAGTCAAGGTCTCCAACAAAGCCCTCCAGCCCCCATCG
 1733 -----+-----+-----+-----+-----+-----+-----+----- 1792
 CCGACTTACCGTCTCATGTTACAGTTCAGAGGTTGTTTCGGGAGGGTTCGGGGGTAGC
 L N G K E Y K C K V S N K A L P A P I E -
 AGAAAACCATCTCAAGCCAAAGGGCAGCCCCGAGAACCACAGGTGTACACCCTGCCCC
 1793 -----+-----+-----+-----+-----+-----+-----+----- 1852
 TCTTTTGGTAGAGTTTCGGTTTCCCGTCGGGGCTCTTGGTGTCCACATGTGGGACGGGG
 K T I S K A K G Q P R E P Q V Y T L P P -
 CATCCCGGGATGAGTGACCAAGAACCAGGTACAGCTGACCTGCCTGGTCAAAGGCTTCT
 1853 -----+-----+-----+-----+-----+-----+-----+----- 1912
 GTAGGGCCCTACTACTGGTTCTTGGTCCAGTTCGGACTGGACGGACCAGTTTCCAAGA
 S R D E T K N Q V S L T C L V K G F Y -
 ATCCCAGCGACATCCCGTGGAGTGGGAGAGCAATGGGCAGCCGGAGAACAACACTACAAGA
 1913 -----+-----+-----+-----+-----+-----+-----+----- 1972
 TAGGGTCGCTGTAGGGCACCTCACCTCTCGTTACCCGTCGGCCTCTTGTGTATGTTCT
 P S D I A V E W E S N G O P E N N Y K T -

14/35

FIG. 4C

CCACGCCTCCCGCTGGACTCCGACGGCTCCTTCTTCCTCTACAGCAAGCTCACCGTGG
1973 -----+-----+-----+-----+-----+-----+----- 2032
GGTGCGGAGGGGACCTGAGGCTGCCGAGGAAGAAGGAGATGTCGTTTCGAGTGGCACC
T P P V L D S D G S F F L Y S K L T V D -

ACAAGAGCAGGTACAGCAGGGGAACGTCTTCTCATGCTCCGTGATGCATGAGGCTCTGC
2033 -----+-----+-----+-----+-----+-----+----- 2092
TGTTCGTCCTGTCGTCCTCCCTTGCGAGAAGAGTACGAGGCACTACGTACTCCGAGACG
K S R W Q Q G N V F S C S V M H E A L H -

ACAACCACTACACAGAAAGAGCCTAAGCTTGTCTGCGGGTAAACCCACCCATGTCAATG
2093 -----+-----+-----+-----+-----+-----+----- 2152
TGTGCGTATGTCTCTTCGATTCTGAACAGACGCCCATTGGGGTGGGTACAGTTAC

IgA tailpiece
N H Y T Q K S L S L S A G K P T H V N V -

Xba I
TGTCTGTTGTCAATGCGGAGGTGGACGGCACCTGCTACTGATAGTCTAGAGCTCGCTGAT
2153 -----+-----+-----+-----+-----+-----+----- 2212
ACAGACAACAGTACGCCTCCACCTGCCGTGGACGATGACTATCAGATCTCGAGCGACTA
S V V M A E V D G T C Y * *

15/35

FIG. 5

Sequence and translation for region encoding the CD86 extracellular domain
in the CD86IgGtp vector

KpnI

TCGGTTACCTGCATATCAAGCTAATTCGGTACCAGCAGAAGCAGCCAAA
657 ----- 707
AGCCAATGGAAGTATAGTTCGATTAAGCCATGGTCGTCTTCGTCGGTTT

ATGGATCCCCAGTCACTATGGGACTGAGTAACATTCTCTTTGTGATGGCCTTCCTGCTC
708 ----- 767
TACCTAGGGGTCTGTGATACCCTGACTCATTGTAAGAGAAACACTACCGGAAGGACGAG

Initiation codon
M D P Q T M G L S N I L F V M A F L L -

TCTGGTGCTGCTCTCTGAAGATTCAAGCTTATTTCAATGAGACTGCAGACCTGCCATGC
768 ----- 827
AGACCACGACGAGAGACTTCTAAGTTCTGAATAAAGTTACTCTGACGTCTGGACGGTACG

Mature Peptide Sequence
S G A A L K I Q A Y F N E T A D L P C -

CAATTTGCAAACCTCAAAAACCAAAGCCTGAGTGAGCTAGTAGTATTTTGGCAGGACCAG
828 ----- 887
GTTAAACGTTTGAGAGTTTGGTTTCGGACTCACTCGATCATCATAAAACCGTCTCTGGTC
Q F A N S Q N Q S L S E L V V F W Q D Q -

GAAAACCTGGTTCGAATGAGGTATACTTAGGCAAAGAGAAATTTGACAGTGTTCATTCC
888 ----- 947
CTTTTGAACCAAGCTTACTCCATATGAATCCGTTTCTCTTTAAACTGTCACAAGTAAGG
E N L V L N E V Y L G K E K F D S V H S -

AAGTATATGGGCTCACAAGTTTGTATTCCGACAGTTGGACCCTGAGACTTCACAATCTT
948 ----- 1007
TTCATATACCCGCTGTGTTCAAAACTAAGCCTGTCAACCTGGGACTCTGAAGTGTTAGAA
K Y M G T S F D S D S W T L R L H N L -

CAGATCAAGGACCTGGGCTTGATCAATGTATCATCCATCACAAAAAGCCCACAGGAATG
1008 ----- 1067
GTCTAGTTCCTGTCCTCGAACATAGTTACATAGTAGGTAGTGTTCGCGGTGTCCTTAC
Q I K D F G L Y Q C I I H H K K P T G M -

16/35

FIG. 5A

ATTCGCATCCACAGATGAATTCTGAACTGTCAGTGCTTGCTAACTTCAGTCAACCTGAA
 1068 ---+-----+-----+-----+-----+-----+-----+----- 1127
 TAAGCGTAGGTCCTTACTTAAAGACTTGACAGTCACGAACGATTGAAGTCAGTTGGACTT
 I R I H M N S E L S V L A N F S Q P E -

ATAGTACCAATTTCTAATATAACAGAAAATGTGTACATAAAATTGACCTGCTCATCTATA
 1128 ---+-----+-----+-----+-----+-----+-----+----- 1187
 TATCATGGTTAAATATTATATTGTCTTTTACACATGTATTTAACTGGACGAGTAGATAT
 I V P I N I T E N V Y I N L T C S S I -

CACGGTTACCCAACCTAAGAAGATGAGTGTTTTGCTAAGAACCAAGAATTCAACTATC
 1188 ---+-----+-----+-----+-----+-----+-----+----- 1247
 GTGCCAATGGGTGGGATTCTTCTACTCACAAAACGATTCTTGGTTCTTAAGTTGATAG
 H G Y P L P K K M S V L L R T K N S T I -

GAGTATGATGGTATATGCAGAAATCTCAAGATAATGTCACAGAAGTGTACGACGTTTCC
 1248 ---+-----+-----+-----+-----+-----+-----+----- 1307
 CTCATACTACCAATATACGTCTTTAGAGTTCTATTACAGTGTCTTGACATGCTGCAAAGG
 E Y D G L M Q K S Q D N V T E L Y D V S -

ATCAGCTTGTCTCTTCATTCCCTGATGTTACGAGCAATATGACCATCTTCTGTATTCTG
 1308 ---+-----+-----+-----+-----+-----+-----+----- 1367
 TAGTCGAACAGACAAAGTAAGGGACTACAATGCTCGTTATACTGGTAGAAGACATAAGAC
 I S L S A S F P D V T S N M T I F C I L -

GAAACTGACAAGATCGGGCTTTTATCTTCACCTTTCTCTATAGAGCTTGAGGACCCCTCAG
 1368 ---+-----+-----+-----+-----+-----+-----+----- 1427
 CTTTGAAGTGTCTTGGCGAAAATAGAAGTGGAAAGAGATATCTCGAACTCCTGGGAGTC
 E T D K T R L L S S P F S I E L E D P Q -

EagI
 CCTCCCCAGACCAAGAGCCCAAATCGGCCGACAAAACCTCACACATGCCACCGTGCCCA
 1428 ---+-----+-----+-----+-----+-----+-----+----- 1487
 GGAGGGGGTCTGCTTCGGGTTTAGCCGGCTGTTTTGAGTGTGTACGGGTGGCACGGGT
 Ig Hinge
 P P P D E E P K S A D K T H T C P P C P -

17/35

FIG. 6

Sequence and translation for region encoding the CMV promoter and CTLA4 extracellular region in the CTLA4Igntp vector.

SPEI

```
GACGTCGACGGGAGATCGGGGATCGATCCGTCGACGTACGACTAGTTATTAATAG
1  -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 60
CTGCAGCTGCCCTCCCTCTAGCCCCCTAGCTAGGCAGCTGCATGCTGATCAATAATTATC

TAATCAATTACGAGTCATTAGTTCATAGCCCATATATGGAGTTCGCGTTACATAACTT
61  -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
ATTAGTTAATGCCAGTAATCAAGTATCGGGTATATACCTCAAGGCGCAATGTATTGAA

ACGGTAAATGGCCCTGGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATG
121 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
TGCCATTACCCGCGGACCGACTGGCGGGTTGCTGGGGGCGGGTAACTGCAGTTATTAC

ACGTATGTTCCCGAGTAACGCCAATAGGGACTTTCATTGACGTCAATGGGTGGACTAT
181 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
TGCATACAAGGGGCATTGCGGTTATCCCTGAAAGGTAAGTGCAGTTACCCACCTGATA

TTACGGTAAACTGCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCT
241 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
AATGCCATTTGAGGTGAACCGTCATGTAGTTCACATAGTATACGGTTCATGCGGGGGA

ATTGACGTCAATGGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTATGG
301 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
TAACTGCAGTTATGCCATTTACCGGGCGGACCGTAATACGGGTCATGTACTGGAATACC

GACTTTCCTACTGACAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGG
361 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
CTGAAAGGATGAGGTGCATGTAGATGCATAATCAGTAGCGATAATGGTACCACCTACGCC

TTTTGGCAGTACCAATGGGCGTGATAGCGGTTTGACTCACGGGGATTTCGAAGTCTC
421 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
AAAACCGTCATGCTTACCCGCACCTATCGCCAAACTGAGTGCCCTAAAGGTTTCAGAG

CACCCCATTGACCAATGGGAGTTTGTGTTGGCGACTCACTATAGGAGTCCCAAGCTT
481 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
GTGGGGTAACTGCAATACCCCTCAAACAAAACCGCTGAGTGATATCCTCAAGGGTTCGAA
```


18/35

FIG. 6A

CTAGAGATCCCTTATGATCCATTGTGCTCTAAAGGACCTGAACACCGCTCCCATAAAGCC
 541 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
 GATCTCTAGGGAATCTAGGTAACACGAGATTTCCTGGACTTGTGGCGAGGGTATTTCGG

ATGGCTTGCCCTTATTTTCAGCGGCACAAGGCTCAGCTGAACCTGGCTGCCAGGACCTGG
 601 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
 TACCGAACGGAAATAAAGTCGCCGTGTTCCGAGTCGACTTGGACCGACGGTCTTGGACC
Initiation sequence
 M A C L F Q R H K A Q L N L A A R T W -

CCCTGCACTCTCTATTTTCTTCTCTTCATCCCTGTCTTCTGCAAAGCAATGCACGTG
 661 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
 GGGACGTGAGACATAAAAAAGAAGAGAAGTAGGGACAGAAGACGTTTCGTTACGTGCAC
Mature
 P C T L F F L L F I P V F C K A M H V -

GCCCAGCCTGCTTCTACTGGCCAGCAGCCGAGGCATCGCCAGCTTTGTGTGTGAGTAT
 721 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
 CGGGTCGGACGAGGCATGACCGGTGCTCGGCTCCGTAGCGGTCGAAACACACACTCATA
Peptide Sequence
 A Q P A V V L A S S R G I A S F V C E Y -

GCATCTCCAGGCAAGGCACTGAGGTCCGGGTGACAGTGCTTCGGCAGGCTGACAGCCAG
 781 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
 CGTAGAGGTCCGGTGGTGACTCCAGGCCCACTGTACGAAGCCGTCCGACTGTCCGGTC

A S P G T E V R V T V L R Q A D S Q -

GTGACTGAAGTCATCCGGCAACCTACATGACGGGGAATGAGTTGACCTTCCTAGATGAT
 841 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
 CACTGACTTCAGATGCGCGTTGGATGTACTGCCCCTTACTCAACTGGAAGGATCTACTA

V T E V C A A T Y M T G N E L T F L D D -

TCCATCTGCACGATACCTCCAGTGGAATCAAGTGAACCTCACTATCCAAGGACTGAGG
 901 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
 AGGTAGACGTGCATGGAGGTACCTTTAGTTCACTTGGAGTGATAGGTTCTGACTCC

S I C T C T S S G N Q V N L T I Q G L R -

19/35

FIG. 6B

GCCATGGACACCTCTCTACATCTGCAAGGTGGAGCTCATGTACCCACCGCCATACTAC
961 -----+-----+-----+-----+-----+
CGGTACCTGTGGGAGATGTAGACGTTCCACCTCGAGTACATGGGTGGCGGTATGATG

A M D T L Y I C K V E L M Y P P P Y Y -

CTGGGCATAGGCGGAACCCAGATTTATGTAATTGATCCAGAACCGTGCCCAGATTCT
1021 -----+-----+-----+-----+-----+
GACCCGTATCCCGTCTTGGGTCTAAATACATTAAGTCTTGGCACGGGTCTAAGA

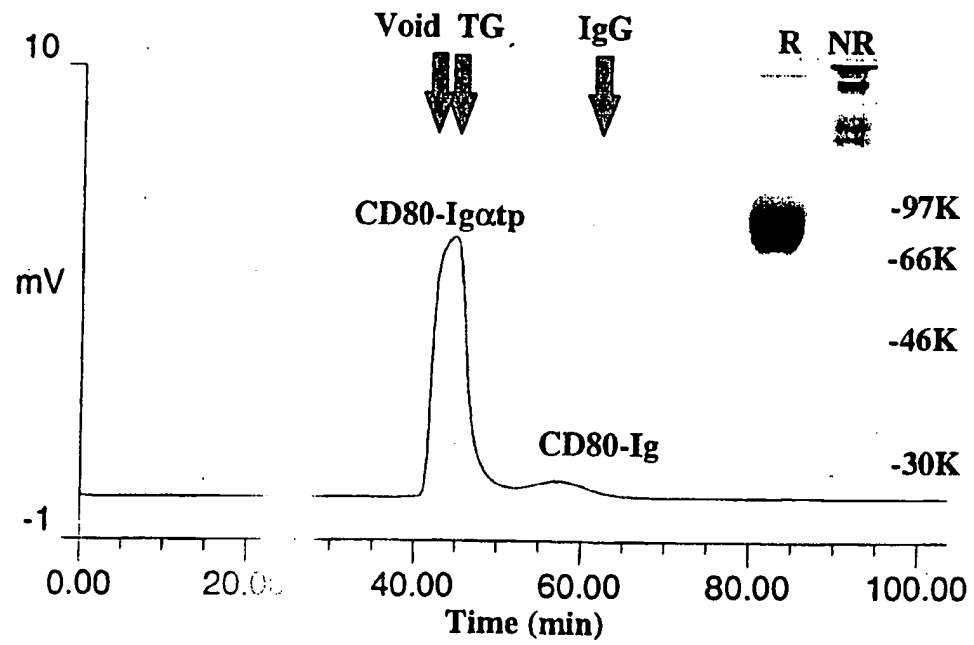
L G I G S T Q I Y V I D P E P C P D S

EAGI
GACGCTGAGCCCTCGGCCG
1081 -----+-----+-----+-----+-----+
CTGCGACTCGGCGAGCCGGC

Ig Hinge
D A E P S A D

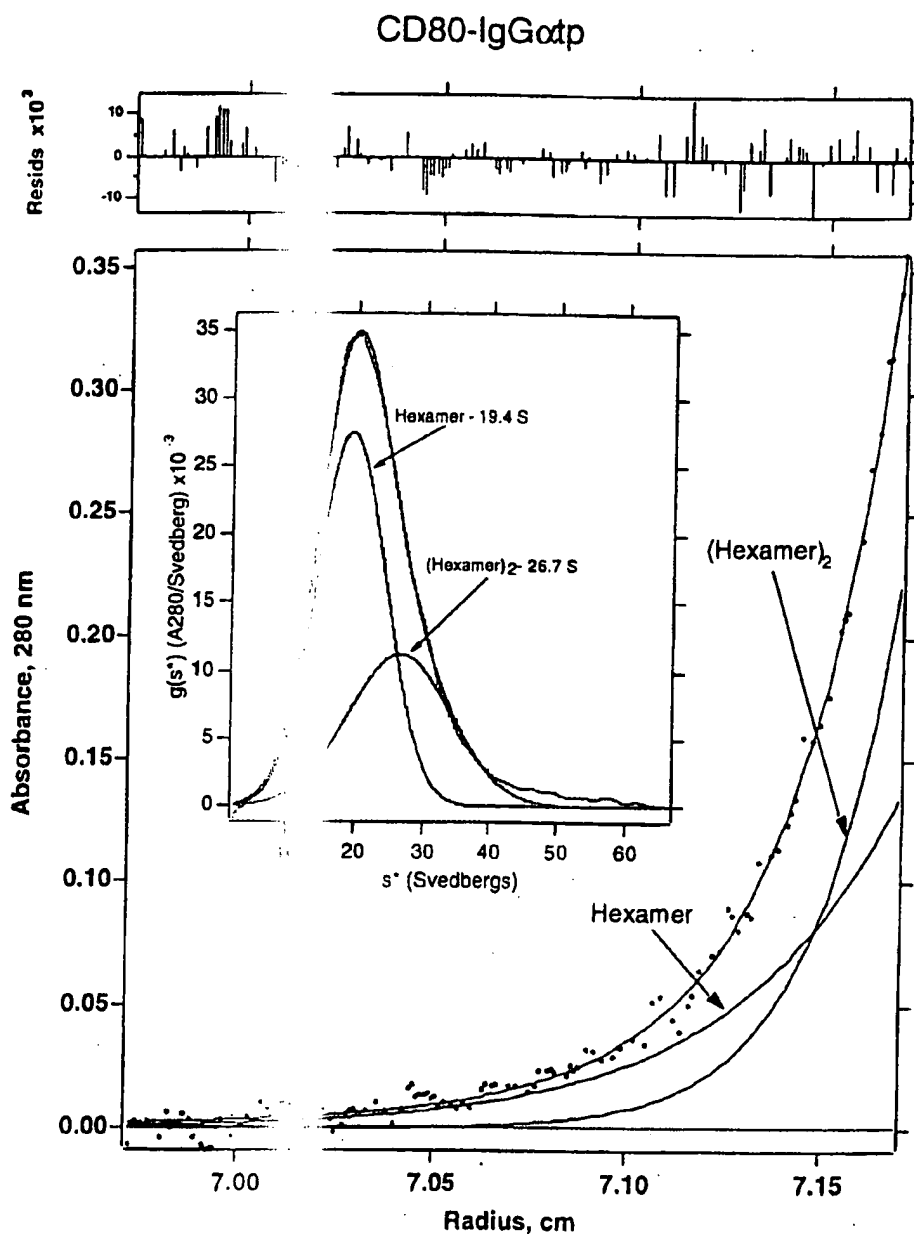
20/35

FIG. 7



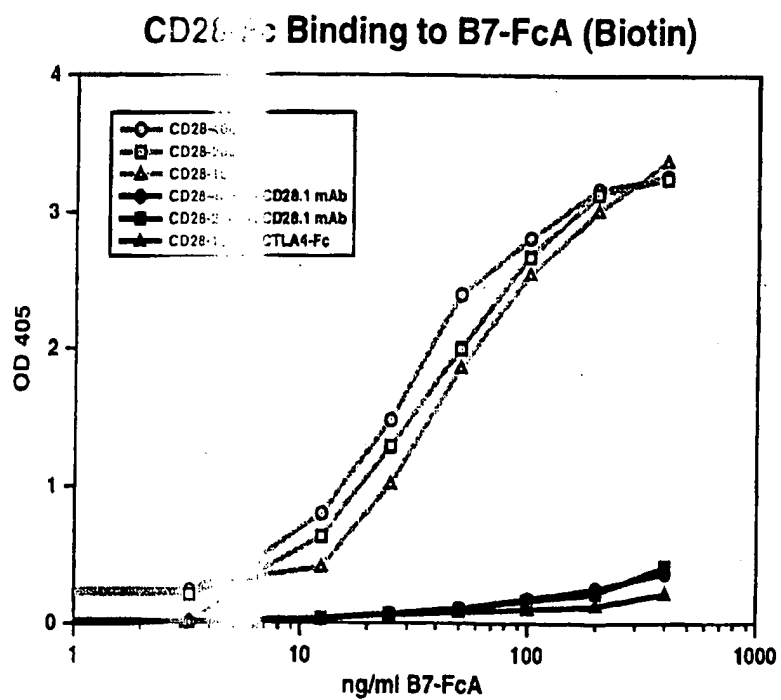
21/35

FIG. 8



22/35

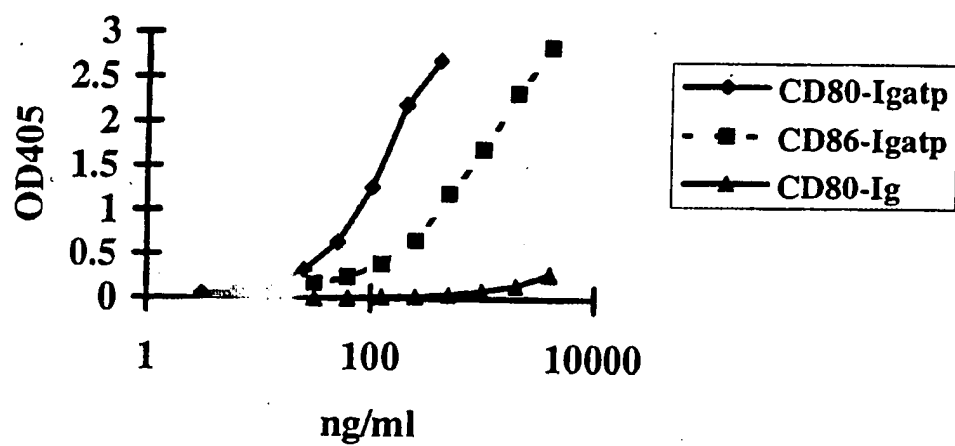
FIG. 9



23/35

FIG. 10

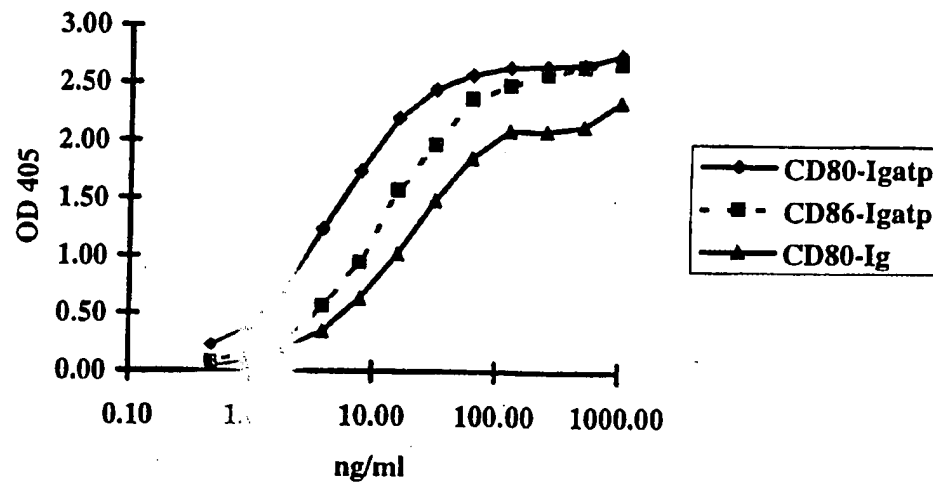
**Saturation binding of CD80, CD86 Fusion
Proteins to CD28-Ig**



24/35

FIG. 11

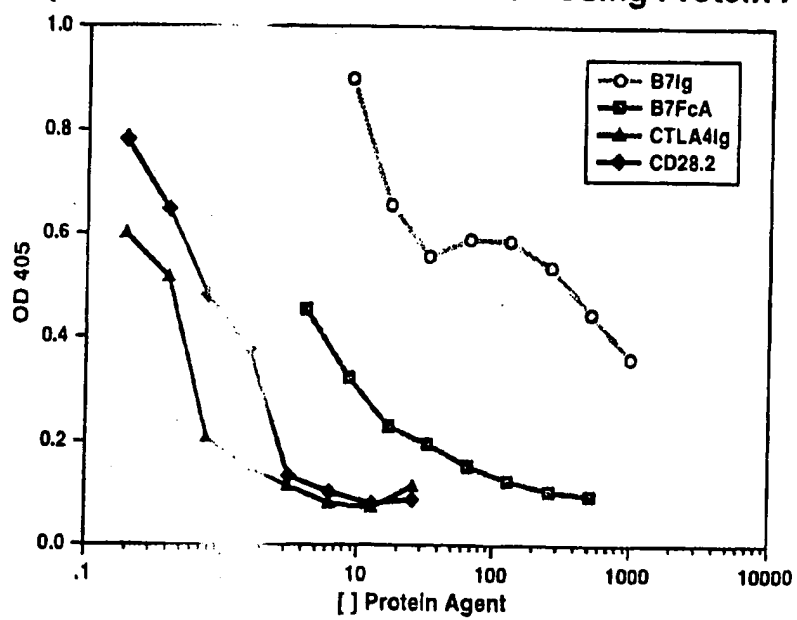
Saturation binding of CD80-Ig α tp, CD86-Ig α tp and CD80-Ig to CTLA4-Ig



25/35

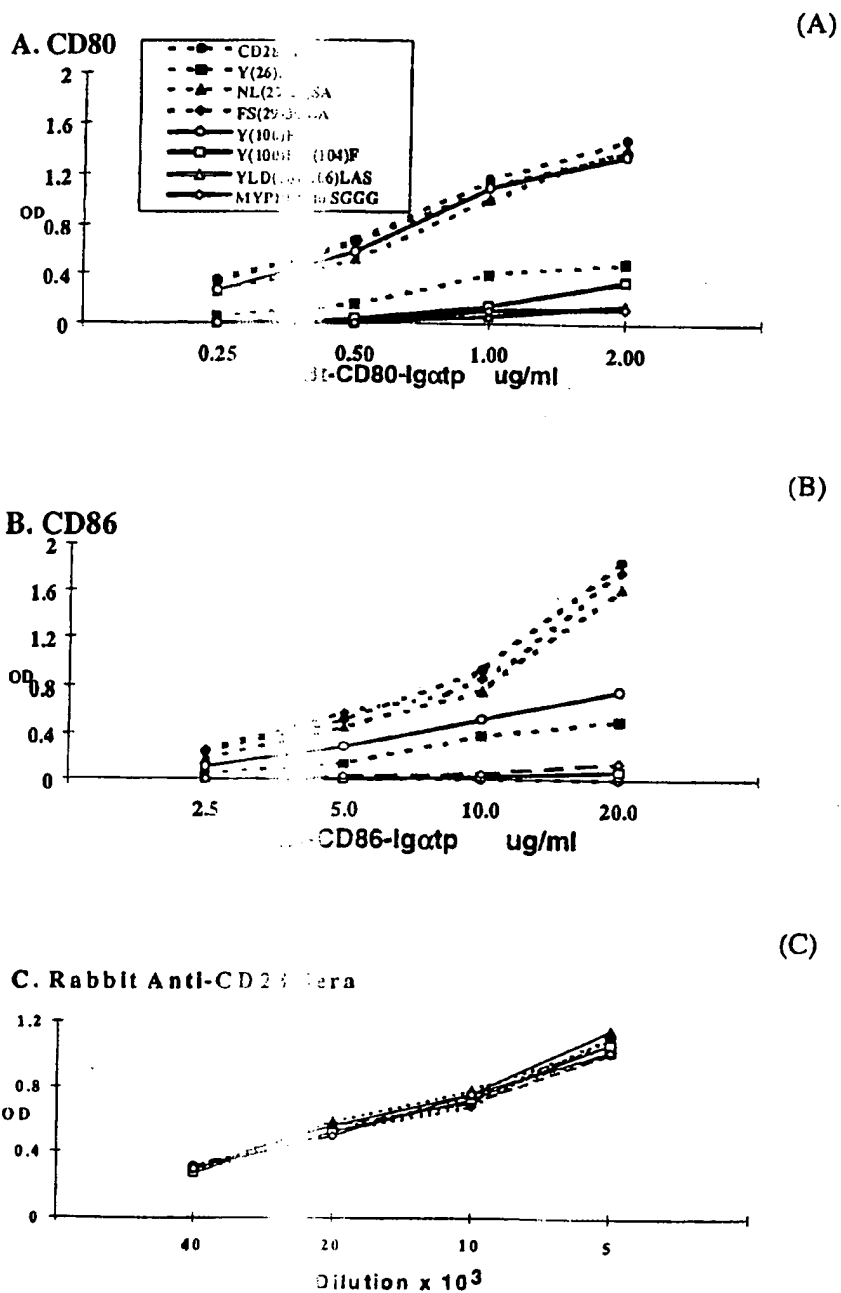
FIG. 12

Competition Curves for CD28/B7FcA Using Protein Agents



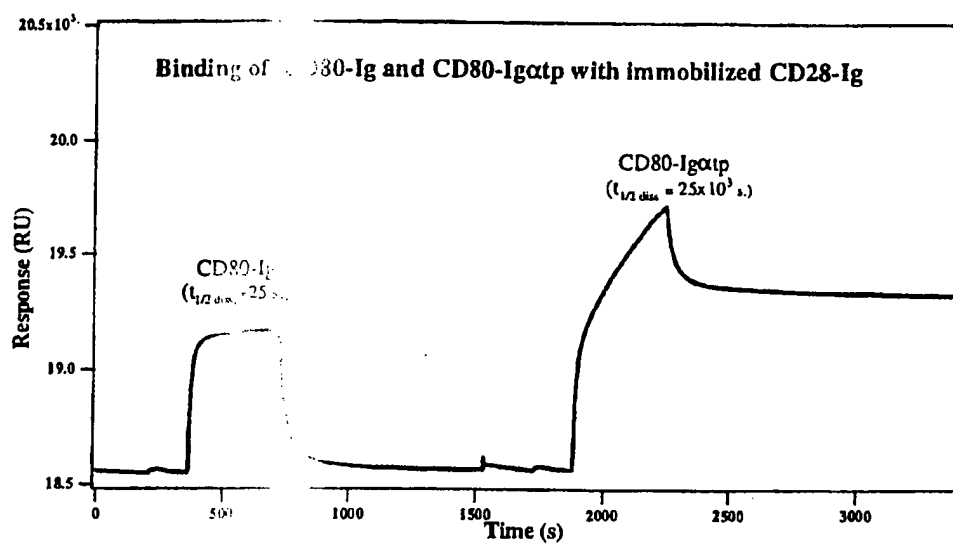
26/35

FIG. 13



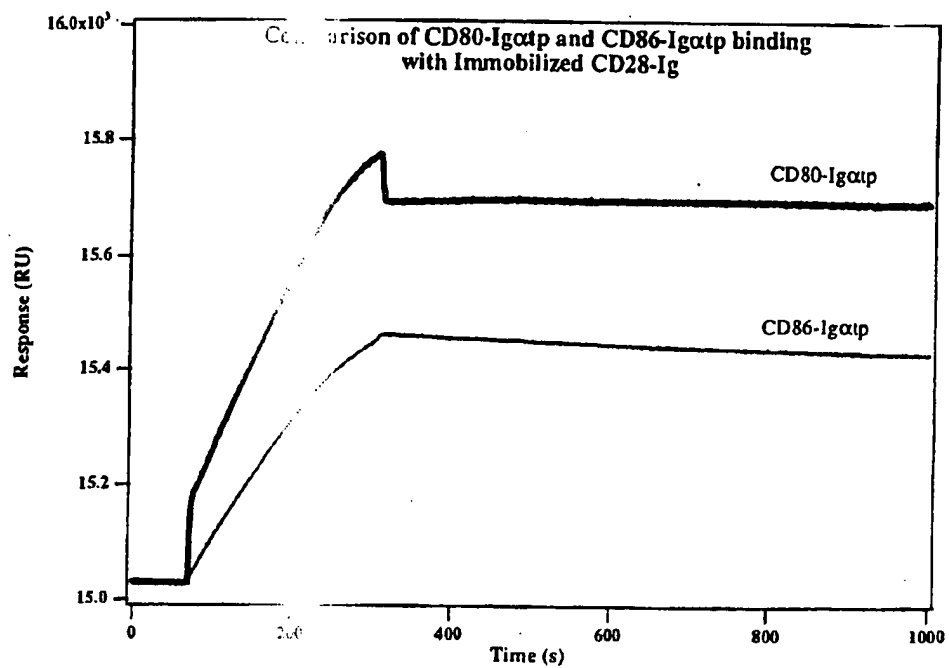
27/35

FIG. 14



28/35

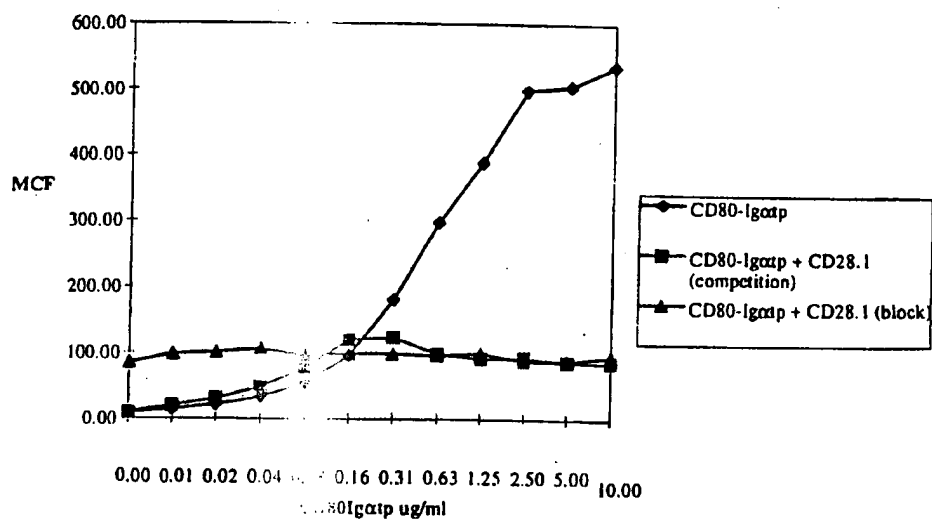
FIG. 15



29/35

FIG. 16

(A)

Flow cytometric analysis of CD80-Ig α p binding to PCD28.1.S2.1 Cells

(B)

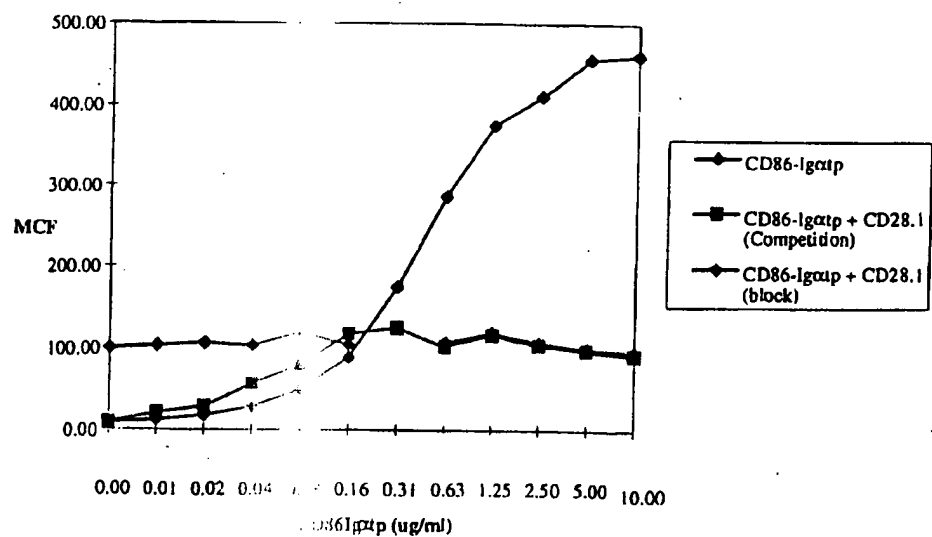
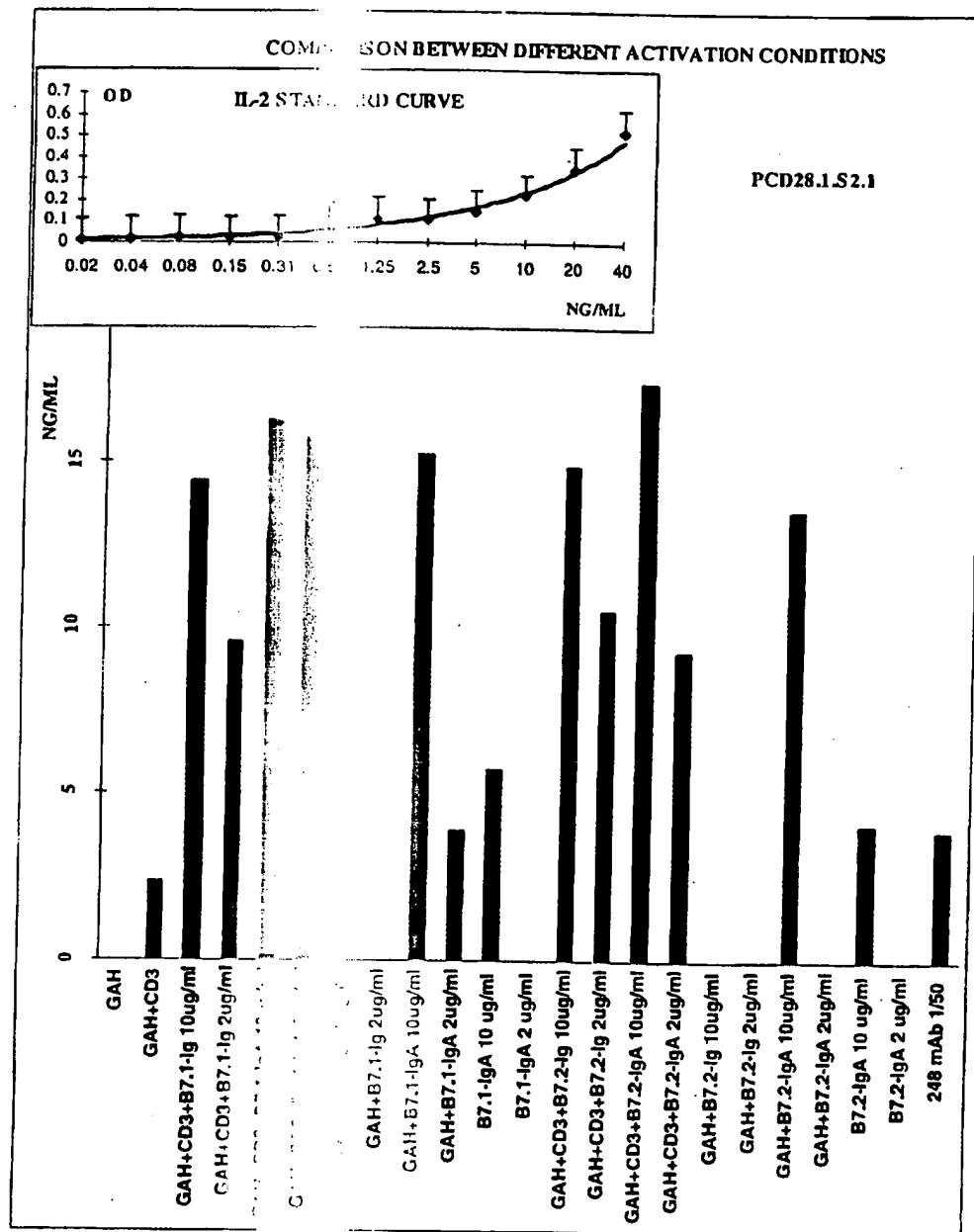
Flow cytometric analysis of CD86-Ig α p binding to PCD28.1.S2.1 Cells

FIG. 17



31/35

FIG. 18

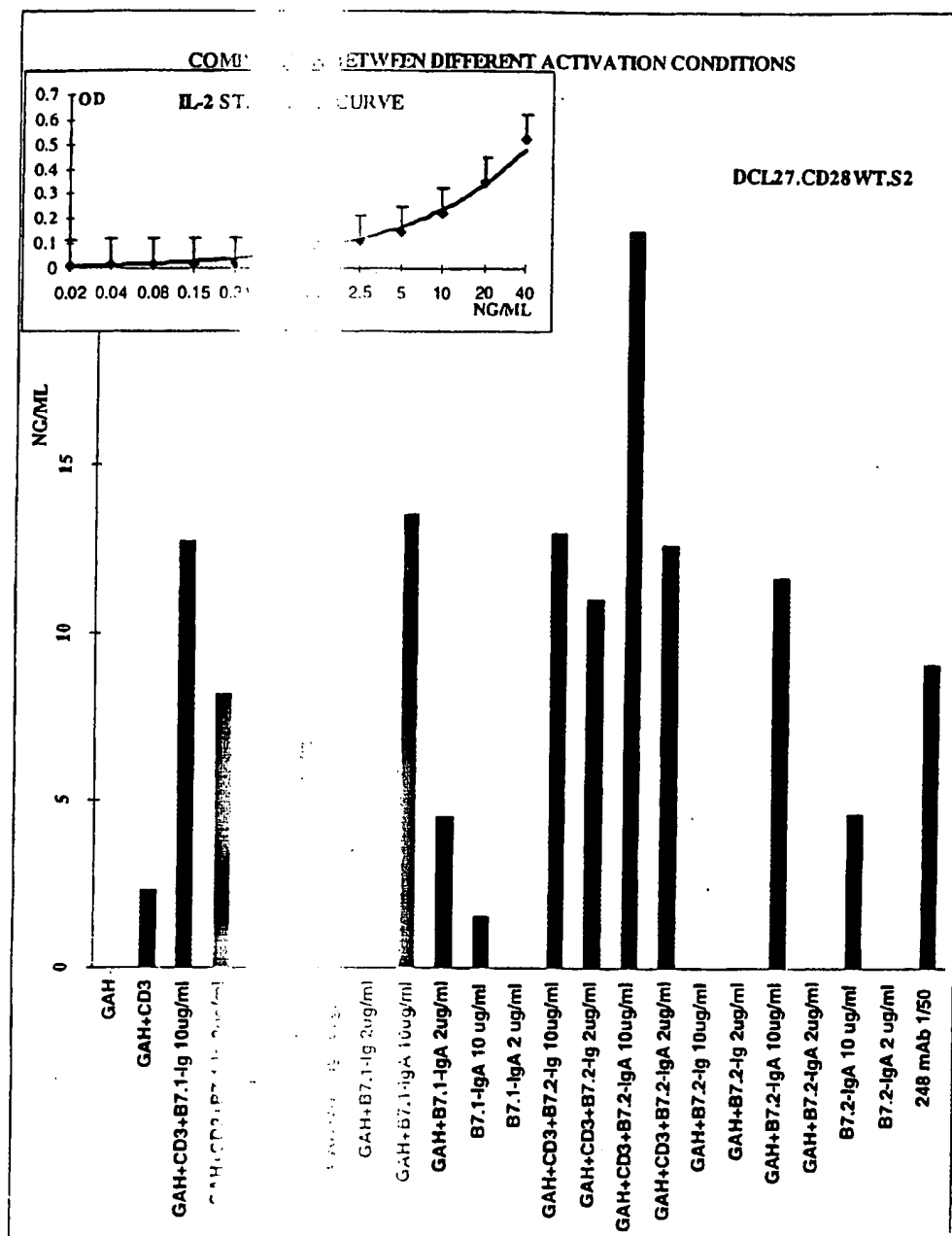
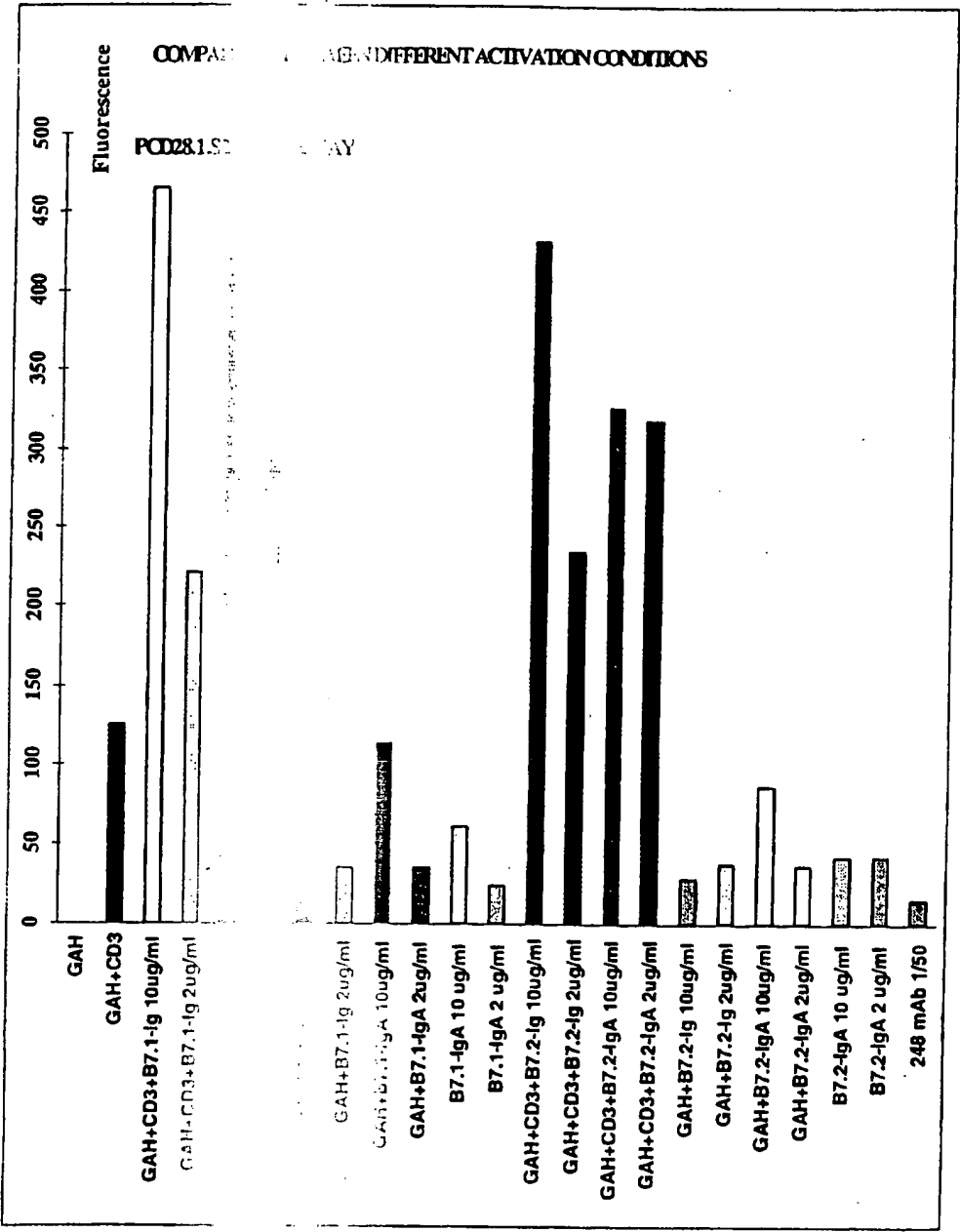
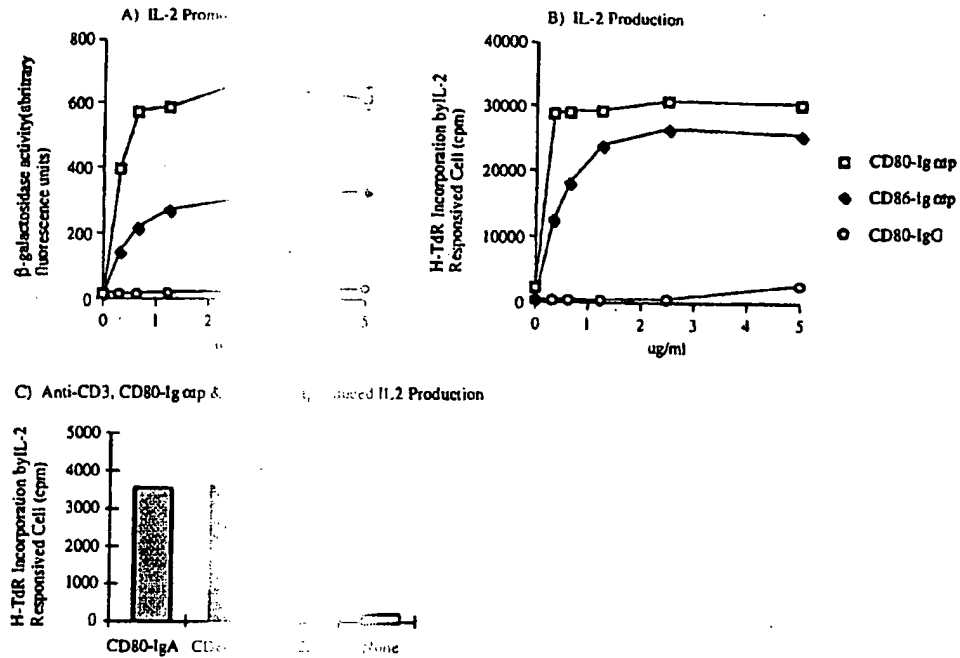


FIG. 19



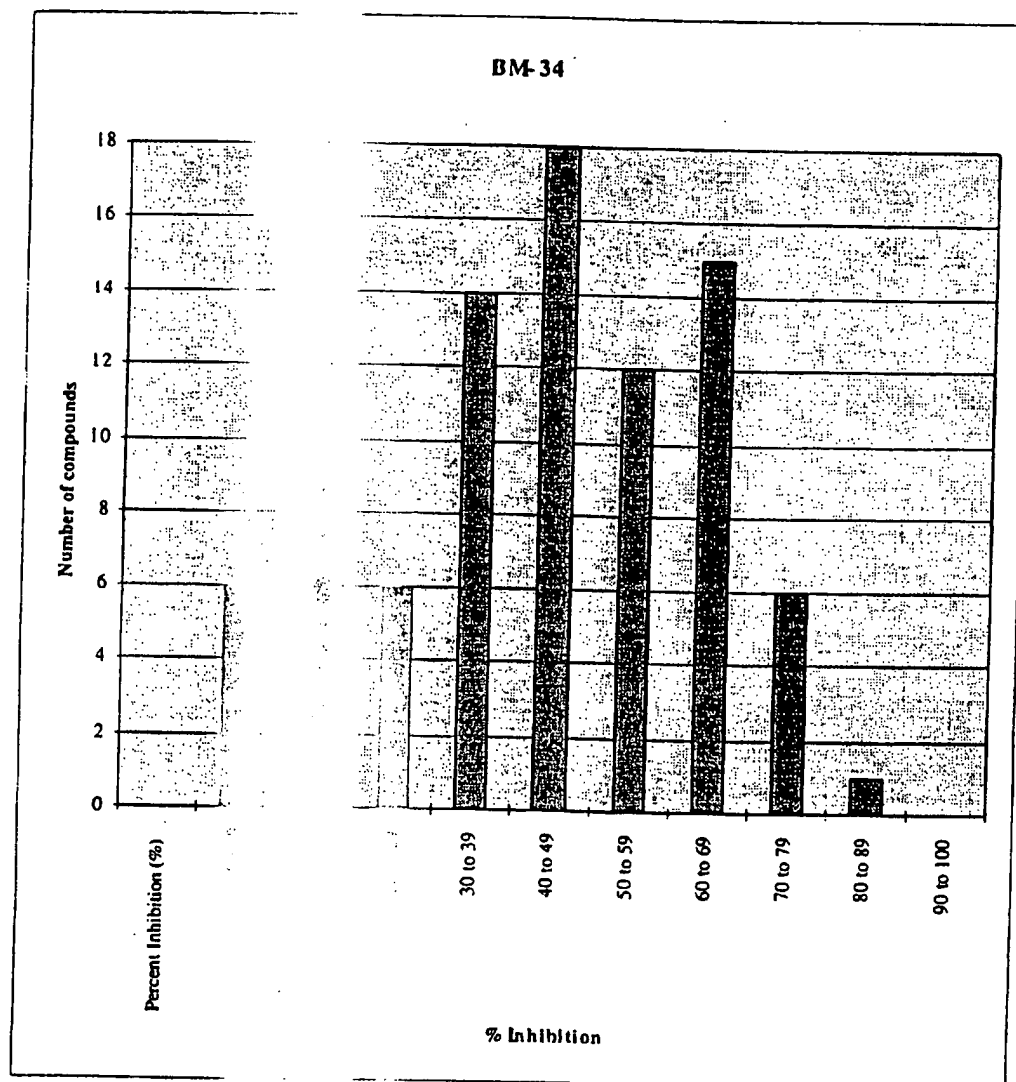
33/35

FIG. 20

CD28-mediated IL-2 Production and IL-2 Production Stimulated by CD80-Ig α p, CD86-Ig α p or anti-CD3

Compounds were incubated with PCD28.1.S2.1 cells and soluble CD80Ig α p or CD86Ig α p (1 ug/ml) or anti-mouse CD3 (1 ng/ml, immobilized) for 24 hours. IL2 in the supernatants was measured by standard assay using an IL2 dependent cell line (CTLL-20). β -galactosidase in the PCD28 cell pellets from 96 well microtiter plates was measured using a fluorogenic substrate.

FIG. 21



35/35

FIG. 22

